MUTATION-SPECIFIC AND COMMON PHOSPHOTYROSINE SIGNATURES OF *KRAS* G12D AND G13D ALLELES

by

Raiha Tahir

A dissertation submitted to The Johns Hopkins University in conformity with the requirement of the degree of Doctor of Philosophy

Baltimore, MD

August 2018

ABSTRACT

KRAS is one of the most frequently mutated genes across all cancer subtypes. Two of the most frequent oncogenic *KRAS* mutations observed in patients result in glycine to aspartic acid substitution at either codon 12 (G12D) or 13 (G13D). Although the biochemical differences between these two predominant mutations are not fully understood, distinct clinical features of the resulting tumors suggest involvement of disparate signaling mechanisms. When we compared the global phosphotyrosine proteomic profiles of isogenic colorectal cancer cell lines bearing either G12D or G13D *KRAS* mutations, we observed both shared as well as unique signaling events induced by the two *KRAS* mutations. Remarkably, while the G12D mutation led to an increase in membrane proximal and adherens junction signaling, the G13D mutation led to activation of signaling molecules such as non-receptor tyrosine kinases, MAPK kinases and regulators of metabolic processes. The importance of one of the cell surface molecules, MPZL1, which found to be hyperphosphorylated in G12D cells, was confirmed by cellular assays as its knockdown led to a decrease in proliferation of G12D but not G13D expressing cells. Overall, our study reveals important signaling differences across two common *KRAS* mutations and highlights the utility of our approach to systematically dissect the subtle differences between related oncogenic mutants and potentially lead to individualized treatments.

Our findings from studies focusing on two different oncogenic mutations of *KRAS* highlighted the common hyperphosphorylation of TNK2, a non-receptor tyrosine kinase that is known to be mutated, overexpressed, or hyperphosphorylated in several cancer subtypes. Non-receptor tyrosine kinases represent an important class of signaling

molecules which are involved in driving diverse cellular pathways. Although the large majority have been well-studied in terms of their protein binding partners, the interactomes of some important non-receptor tyrosine kinases such as TNK2 (also known as activated Cdc42-associated kinase 1 or ACK1) have not been systematically investigated. Aberrant expression and hyperphosphorylation of TNK2 have been implicated in a number of cancers, although the exact proteins and cellular events that mediate phenotypic changes downstream of TNK2 are unclear. Biological systems that employ proximity-dependent protein labeling methods, such as biotinylation identification (BioID), are being increasingly used to map protein-protein interactomes as they provide increased sensitivity in finding interacting proteins. Therefore, as an extension of our effort to understand oncogenic *KRAS* signaling and its components, we sought to map the interacting partners of TNK2. We also employed Stable Isotope Amino Acid Labeling in Cell Culture (SILAC) to quantitatively explore the interactome of TNK2. By performing a controlled comparative analysis between full-length TNK2 and its truncated counterpart, we were not only able confidently identify site-level biotinylation of previously well-established TNK2 binders and substrates (NCK1, NCK2, CTTN, STAT3), but also identify of several novel TNK2 interactors. We validated TNK2 interaction with one novel interactors, clathrin interactor 1 (CLINT1), using immunoblot analysis. Overall, this work highlights several molecules that warrant further experiments to assess their functional significance in TNK2-mediated signaling.

DISSERTATION REFEREES

Graduate Advisor: **Akhilesh Pandey, M.D. Ph.D.,** Professor, McKusick-Nathans Institute of Genetic Medicine, Departments of Biological Chemistry, Oncology, and Pathology

Thesis Committee Member: **Michael Wolfgang, Ph.D.,** Professor, Department of Biological Chemistry, and Center for Metabolism and Obesity Research

Thesis Committee Member: **Edward Gabrielson, M.D.,** Professor, Departments of Oncology, and Pathology

ACKNOWLEDGEMENTS

I would like to thank several people that made it possible for me to compose this dissertation and perform the work described within. I would like to thank the Biochemical, Cellular, and Molecular Biology Program (BCMB) at the Johns Hopkins School of Medicine for investing their resources in order to provide me with the best possible scientific training. I would like to especially thank my graduate mentor, Akhilesh Pandey, for agreeing to mentor me in his lab, and for helping me improve by ability to think and solve problems. I would like to thank the members of my thesis committee, Drs. Edward Gabrielson, Heng Zhu, and Michael Wolfgang, who have done a great job overseeing and guiding the trajectory of my thesis work. I am very thankful for companionship of the past and present Pandey lab members, including Jevon Cutler, Chan-Hyun Na, Xinyan Wu, Min-Sik Kim, Chris Mitchell, Daein Kim, Saddiq Zahari, and others. I cannot forget to thank my undergraduate research co-mentors Dr. Sarah Elsea (now at Baylor College of Medicine) and Dr. Amanda Dickinson (Virginia Commonwealth University) for helping me realize my passion for science as well as laying a solid foundation for my scientific training. I also want to acknowledge all the graduate students I have worked with over the years who have helped me acquire and fine tune my technical skills in the lab. Finally, I want to thank my wonderful parents for showering me with their love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABL	Abelson tyrosine kinase
ABI1	Abl interactor 1
ALDH3A1	Aldehyde Dehydrogenase 3 Family Member A1
BioSITe	Biotin Site Identification Technology
CALM	calmodulin
CD73	5'-nucleotidase ecto
CLTC	clathrin heavy chain 1
CLINT1	Clathrin interactor 1
CRKL	CRK-like proto-oncogene adaptor protein
CTTN	cortactin
DCBLD2	discoidin, CUB, and LCCL domain containing 2
EGFR	epidermal growth factor receptor
FYN	tyrosine-protein kinase Fyn
GAK	Cyclin G associated kinase
GAPs	GTPase-activating proteins
GEF	guanine nucleotide exchange factors
GRB2	growth factor receptor-bound protein 2
HRAS	Harvey rat sarcoma viral oncogene homolog
HSP	heat shock protein
HSPB1	heat shock protein beta-1
ITIM	Immunoreceptor Tyrosine-based Inhibitory Motif
KRAP	KRAS-induced actin-interacting protein

KRAS	Kirsten rat sarcoma 2 viral oncogene homolog
LC-MS/MS	liquid chromatography and tandem mass spectrometry
LGALS1	galectin-1
LMO7	LIM domain only protein 7
LYN	tyrosine protein kinase Lyn
MET	hepatocyte growth factor receptor
MPZL1	myelin protein zero-like 1
NCK1	NCK adaptor protein 1
NCK2	NCK adaptor protein 2
NRAS	neuroblastoma RAS viral oncogene homolog
PTM	post translational modification
PPI	protein-protein interaction
pTyr	phosphotyrosine
PTPN11	Tyrosine-protein phosphatase non-receptor type 11
RTK	receptor tyrosine kinase
siRNA	small interfering RNA
STAT	signal transducer and activator of transcription
SOS1	son of sevenless homolog 1
TNK2	tyrosine kinase non-receptor 2
TTI1	TELO2-interactor 1

CHAPTER 1.

INTRODUCTION

Cancer is the second leading cause of death worldwide (1). At its basis, cancer is a genetic disease largely originating from somatic alterations, such as a missense mutations and genomic rearrangements. Such alterations that drive cancer often occur in genomic regions that code for proto-oncogenes and tumor suppressor genes. These alterations result in one or more genes that code for an aberrant protein product that has an increased or decreased biochemical activity compared to the product of the wild-type genes. Genomic landscape sequencing studies across several cancers have found mutations in RAS genes to be the most prevalent feature in all cancer subtypes (2). The RAS superfamily is comprised of more than 100 genes that code for small GTP-binding proteins. Among this family are three prototypical RAS proteins that are the most relevant when discussing cancer mutations: *HRAS* (Harvey rat sarcoma viral oncogene homolog), *KRAS* (Kirsten rat sarcoma 2 viral oncogene homolog), and *NRAS* (neuroblastoma RAS viral oncogene homolog) (3). These genes have ~88% sequence homology. They differ in their C terminal end sequence known as the hypervariable region. Alternate splicing of KRAS within exon 4 can result in the expression of two different splice variants, KRAS 4A and KRAS 4B, though KRAS 4B is the dominant isoform.

While the pattern of mutational frequency for RAS genes varies across cancer subtypes, *KRAS* is the most frequently mutated across all cancers. Cancer subtypes where KRAS mutations are most common include pancreatic ductal adenocarcinoma (~98%), colorectal cancer (~45%), and non-small cell lung cancer (~20%) (4,5). In addition to differences in mutational frequency across cancers, variation is also observed in the frequency of KRAS mutations at specific codons, and the particular substitution mutation (2,6). Mutations at codon 12 account for close to 90% of KRAS mutated cancers.

Oncogenic KRAS mutations, especially in colorectal cancer tumors, largely result in a glycine to aspartic acid substitution at codons 12 (G12D) or 13 (G13D) **(Figure 1.1A)**.

To comprehend the basis of the KRAS-driven biological processes, we need to understand its canonical biochemical activity. The KRAS gene codes for a small membrane-bound GTPase protein that switches structural conformations corresponding to an active and inactive state through the binding and hydrolysis of the nucleotide GTP, respectively **(Figure 1.1B)**. While this catalytic activity is intrinsic to KRAS itself, the rate of GTP exchange and hydrolysis is often regulated through the binding of guanosine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs). GEFs increase the rate at which KRAS dissociates from GDP in exchange for GTP, while GAPs stimulate the intrinsic ability of KRAS to hydrolyze GTP (7). The dynamic binding to GAPs/GEFs is orchestrated by upstream signals, such as those from ligand-induced activation of receptor tyrosine kinases. In this manner, KRAS serves as a highly regulated molecular switch for mediating signal transduction. Activation of KRAS also results in a structural conformation that favors protein-protein interaction with other effector binding partners of KRAS that allow for diversity in the signaling axes activated downstream of KRAS. In addition to the regulation of its catalytic activity, many enzymes are also involved to ensure the proper localization of KRAS at the plasma membrane. To form a fully functioning protein, KRAS requires processing via several post-translational modifications. Like the C-terminal end of most Ras proteins, the last four residues of the protein sequence of KRAS harbors a CAAX box that is severs as a recognition motif for farnesyltransferases, which catalyze

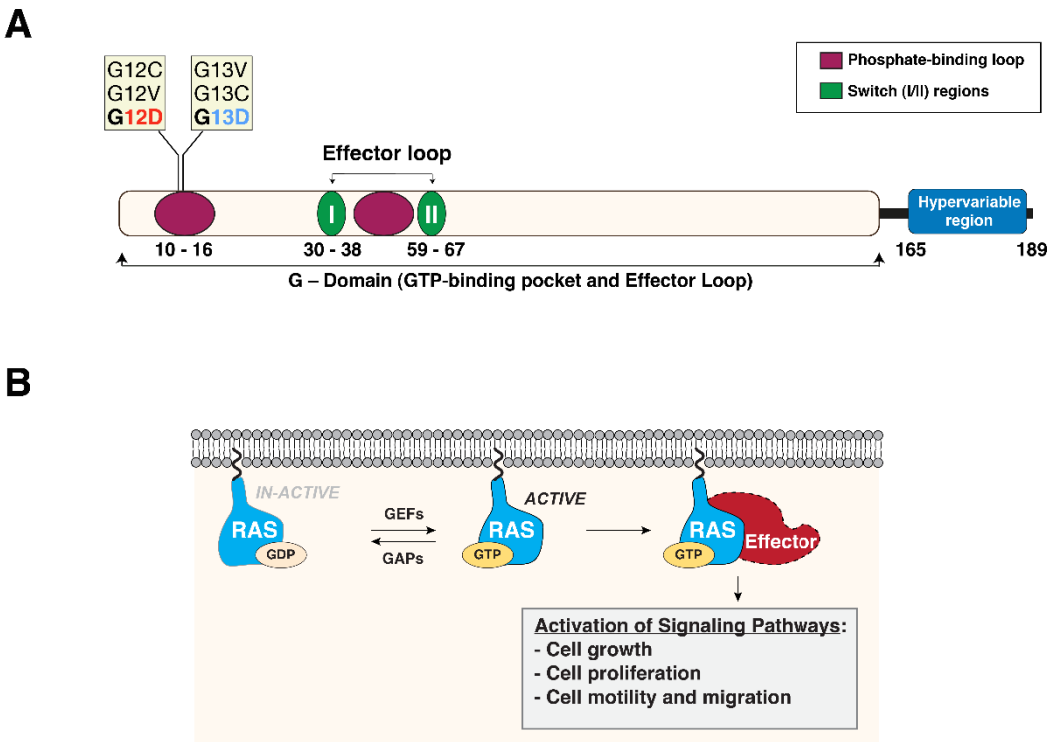

Figure 1.1 Distinct phenotypes observed in codon 12 mutations in KRAS. (A) Schematic depicting the domain architecture of KRAS protein along with the location of two mutational hotspots. Various domains are indicated as colored ovals and the hypervariable region is shown in blue at the C-terminus. Amino acid positions of the domains are marked. (B) Schematic summarizing the alteration in activation state of RAS in its GDP and GTP bound state. One of the cellular events following RAS activation, namely the physical association with downstream effector proteins and activation of downstream pathways, is also depicted.

the transfer and attachment of a farnesyl lipid to cysteine residues. Farnesylation of KRAS is crucial for mediating its binding and localization to endomembranes such as the surface of the endoplasmic reticulum. In addition to farnesylation, other C-terminal features and modifications that are important for membrane localization of KRAS include the C-terminal lysine rich region, carboxymethylation of this cysteine by Isoprenylcysteine carboxylmethyltransferase (ICMT1), and proteolytic cleavage of the terminal residues by CAAX prenyl protease 2 (RCE1).

The ability of substitution mutations of *KRAS* to promote oncogenicity lies in their ability to cause aberrant GTPase cycling, which renders the protein in a prolonged active state. Several hypotheses have been proposed regarding the exact mechanism behind the hyperactivation of mutant KRAS. Since many oncogenic mutations of KRAS are in its nucleotide binding pocket, including mutations at codons 12, 13, 61, 117, and 146, it is hypothesized that mutations interfere with the protein's ability to stabilize the nucleotide bound to the active site (8,9). Mutations have also been shown to modulate the binding affinity of KRAS to GAPs, thus hindering GAP-induced GTPase activity and inactivation of KRAS (8). Consequently, these changes to the biochemical activity of KRAS cause constitutive activation of KRAS-induced pathways that promote cell proliferation and metastasis.

The regulated activation of KRAS is important for the context-specific activation of several downstream pathways. Activated KRAS at the cell surface is capable of physically associating with key early signaling proteins, especially kinases and adaptor proteins, that initiate various signaling axes that are responsible for promoting cell growth and proliferation. The most notable of these pathways consists of the RAF/MEK/MAPK

axis. Activated KRAS binds and contributes to the activation of RAF kinase, which subsequently leads to the activation of MAPK (mitogen-activated protein kinase) kinases, including MEK (MAP kinase kinase 1) and ERK1/2 (Extracellular regulated kinase 1/2). Activation of KRAS also results in signal transduction through the PI3K/AKT pathways (3). In this pathway, the binding of the p110 catalytic subunit of phosphoinositide-3-kinases (PI3Ks) to activated KRAS results in activation of signaling through pathways involving the Ser/Thr kinase AKT or PKB, and the transcription factor NF-kB (nuclear factor kB). Other important signaling pathways downstream of KRAS are activated by binding to Ral guanine nucleotide exchange factors (RalGEFs), phospholipase-C (PLC), Ras And Rab Interactor 1 (RIN1), T Cell Lymphoma Invasion And Metastasis 1 (TIAM1), Ras association domain-containing protein (RASSF), and other effectors. It is important to note that the specific pathway activated downstream of KRAS is determined by several factors, including the cell type, the upstream stimulus (growth factors, cytokines, etc), regulation by diverse GAPs/GEFs, the binding affinity and specificity of KRAS to its downstream effectors, and other context specific factors.

Mutant forms of *KRAS* in cancer has been found to drive oncogenic transformation through a wide variety of signaling mechanisms. In addition to hyperactivation of the canonical pathways induced as part of normal KRAS-mediated signaling, other signaling strategies are also employed that sometimes serve as a mechanism to bypass normal regulatory pathways. Studies in model systems either harboring endogenous *KRAS* mutations or overexpressing a mutant form of KRAS have identified many such signaling mechanisms that are associated with oncogenic *KRAS*-driven cancers. Various functional studies have characterized the importance of Yes1 activated protein 1 (YAP1) and β-

catenin (CTNNB1) for signaling downstream of *KRAS* mutations. YAP1 and CTNNB1 have shown to be overexpressed in pancreatic ductal adenocarcinomas harboring a *KRAS* mutation (10–12). In another set of studies, calmodulin (CaM) was shown to play a crucial role in oncogenic KRAS-mediated signaling. By directly binding to KRAS and PI3K, calmodulin allows mutant KRAS to activate PI3K/AKT proliferative pathways even in the absence of upstream signals from receptor tyrosine kinases (13,14). Finally, another important oncogenic mechanism employed by mutant *KRAS* involves metabolic reprogramming with cancer cells and tumors. A plethora of studies have identified mutant *KRAS*-induced rewiring of diverse metabolic pathways (15–19). Several key metabolic changes associated with oncogenic KRAS include increased used of glycolytic pathways (increased glucose uptake and sensitivity to glucose deprivation) (20), reprogramming of glutamine metabolism (21), increased used of autophagy pathways (22,23), and increased oxidation of fatty acids (24).

The knowledge derived from decades of studies on the biochemical and molecular workings of KRAS has been applied and leveraged to the previous and still ongoing efforts to develop therapeutics against *KRAS*-mutated tumors (25). The identification of small molecules that selectively bind and inactivate mutant KRAS is a challenging problem, forcing studies to search for alternative strategies to indirectly target mutant KRAS (26,27). Initial therapeutic efforts were directed towards developing inhibitors of enzymes involved in regulating KRAS membrane interaction and subcellular localization. For instance, competitive farnesyltransferase inhibitors (FTIs) as well as inhibitors against KRAS processing enzymes RCE1 and ICMT1, were developed and tested for their ability to disrupt *KRAS*-driven transformation (28,29). However, the antagonism of tumor growth

observed through FTIs was later found to be due to inhibition of other cellular CAAX-motif harboring proteins that are substrates of these enzymes. The other main strategy has been to pharmacologically target signaling molecules found to be essential for sustaining the viability of cells and tumors harboring *KRAS* mutations (30–33). Recent efforts have led to the development of a small molecule, ARS853, that can selectively bind and inactivate the G12C mutated form of KRAS protein (8,34,35). Many of the recent efforts to target mutant KRAS tumors have been informed by a collection of studies that have attempted to map the genes whose knockdown or perturbation is synthetically lethal in the presence of mutant KRAS. Such studies have employed pools of siRNAs/shRNAs (36–39) and more recently CRISPR/Cas9 technologies to screen for such candidates (37,40–42). Several notable genes that have been highlighted and later shown to have promising therapeutic potential include Polo-like kinase 1 (PLK1) (39,43), TANK-binding kinase 1 (TBK1) (44), Enhancer of rudimentary homolog (ERH) (45), and Serine/threonine-protein kinase 33 (STK33) (46). A recently developed small molecule that modulates protein-protein interactions between PLKs was recently found to have a therapeutic effect in mutant KRAS-expressing tumors (47).

Several phenotypic characteristics have been attributed to oncogenic KRAS-mediated transformation and tumor growth. A key property shared by KRAS-mutated tumors is the ability to show resistance to EGFR inhibitors (48–53). Data acquired from clinical trials and studies has significantly contributed to our understanding of KRAS mutant cancers. Inhibition of EGFR was considered a potentially promising strategy to treat colorectal cancer, as it is highly expressed in these tumors. However, retrospective analysis of clinical data on cancer patient's response to treatment with EGFR inhibitors

revealed a strong correlation between presence of a KRAS mutation and lack of response to these inhibitors. Since the reporting of these findings, many studies have attempted to investigate the mechanism of this resistance. The genomic amplification or occurrence of activating mutations in other proto-oncogenes in the presence of mutant *KRAS*, especially the genomic region coding for Hepatocyte growth factor receptor (*MET)*, have been correlated with increased resistance to clinical inhibitors against EGFR (54–56). However, the molecular events and mechanisms that allow *KRAS* to evade the antagonism to EGFR and other RTKs are still unclear.

Even after decades of studies on biochemical characteristics and functional context of KRAS in normal and oncogenic signaling, observations in recent years have revealed additional layers of complexity regarding KRAS mutations and tumorigenesis continue to unravel. Furthermore, the importance of distinguishing between individual hotspot mutations in KRAS is being increasingly appreciated due to accumulating in vitro and clinical studies that report observing differences in phenotypic and clinical behavior between tumors harboring mutations at codon 12 or codon 13 (57–62). Differences in anchorage-independent growth, apoptotic tendency, and metastatic efficiency have been reported (63–66). Codon 12 mutations tend to form spheroid shaped colonies that grow in size and density over time, while codon 13 mutants form smaller colonies. Codon 12 mutants also exhibit increased resistance/protection from apoptosis relative to codon 13 mutants which show higher sensitivity to confluency-dependent apoptosis. In terms of clinical presentation, colorectal cancer patients with tumors harboring codon 12 mutations show a worse overall survival and prognosis as compared to codon 13 mutated tumors (67,68). Furthermore, codon 12 and 13 also exhibit differential degrees of resistant to

EGFR inhibitors, with recent clinical studies reporting that G13D mutant patient tumors are significantly more responsive to the anti-EGFR monoclonal antibody cetuximab compared to codon 12 mutated patient cohorts (58,62).

The molecular basis of the phenotypic and clinical differences between tumors harboring different oncogenic substitution mutations of *KRAS* is still unclear. The lack of understanding regarding the molecular differences in KRAS mutant-specific signaling is, in part, owing to a scarcity of studies utilizing controlled biological systems that allow direct comparison of individual KRAS mutations. Findings from a heterogeneous mutational background that is characteristic of patient samples complicate the interpretation of signaling events contributed largely by the activity of mutant KRAS. These complications can be avoided by using an isogenic system for different KRAS mutants. Several studies have leveraged isogenic systems to study differential signaling between wild-type KRAS and mutant KRAS (69–76). However, fewer studies have used isogenic systems to explore mutation-specific KRAS signaling (66,70,77). One such system that was used previously by Prior et al. profiled the serine and threonine phosphoproteome of isogenic SW48 cells harboring either KRAS mutations. This work identified proteins and phosphorylation sites showing increased abundance specifically in G13D or codon 12 mutants (77). However, the role of tyrosine phosphorylation-mediated signaling in oncogenic *KRAS*-driven cancers has largely been unstudied, as a majority of these signaling events are perceived as upstream of KRAS and thus irrelevant to mutant KRAS induced signaling events downstream. Although tyrosine phosphorylation accounts for only ~1% of total phosphorylation, tyrosine kinases play a disproportionately large role in diseases, especially in cancer. Our previous study revealed that activation of serine and

threonine kinases by the PI3K-AKT signaling pathway lead to widespread modulation of the tyrosine phosphoproteome (78,79). In the present study, we examined mutation-specific KRAS signaling by characterizing the phosphotyrosine proteome along with total proteome downstream of G12D and G13D mutations in the same cell background. Our analysis provided insights into changes in phosphotyrosine and total proteome that were driven by both KRAS mutants, as well as changes that were uniquely induced by each mutant.

CHAPTER 2.

MATERIALS AND METHODS

2.1 Plasmids, cloning, antibodies, and reagents – Isogenic SW48 cells used in this study were purchased from Horizon Discovery. Cells with heterozygous knock-in of oncogenic KRAS mutation G12D (cat. no. HD 103-011) and G13D (cat. no. HD 103-002) were used along with the KRAS wild-type parental cells (cat. no. HD PAR-006). All cells were cultured in RPMI 1640 supplemented with fetal bovine serum (FBS), L-glutamine, and antibiotics, which were purchased from Invitrogen (Carlsbad, CA). RPMI 1640 SILAC media deficient in L-lysine and L-arginine were from Thermo Fisher Scientific (Waltham, MA). SILAC amino acids; $^{2}H_4$-Lysine (K4); $^{13}C_6$, $^{15}N_2$-Lysine (K8); $^{13}C_6$-Arginine (R6) and $^{13}C_6$, $^{15}N_4$-Arginine (R10); were from Cambridge Isotope Laboratories (Andover, MA). TPCK-treated trypsin was from Worthington Biochemical Corp. (Lakewood, NJ). Anti-phosphotyrosine mouse mAb (P-Tyr-1000) beads for immunoaffinity purification of phosphopeptides were from Cell Signaling Technology (Danvers, MA). The primary antibodies used in this study were as follows: Antibodies against LGALS1 (12936), PZR (9893), and phospho-PZR Y263 (8088) were purchased from Cell Signaling Technology (Danvers, MA). Antibodies against ALDH3A1, anti-phosphotyrosine (clone 4G10) were purchased from Millipore (Belerica, MA). Antibodies against KRAS (clone F235) and TNK2 (clone A-11) were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Anti-biotin antibody (Bethyl Laboratories, #150-109A), streptavidin-HRP (Abcam, #ab7403), protein G beads (EMD Millipore, #16-266), biotin (Sigma Aldrich, #B4501), lipofectamine 2000 (Thermo Fisher Scientific, #11668019). The plasmid pcDNA3.1 BioID-HA containing the mutant BirA-R118G was purchased through Addgene (Cambridge, MA, USA) (Plasmid #36047). The BioID-HA cassette from above plasmid

was used to generate pBABE BioID-HA plasmid. The full-length (FL) or N-terminal half (ΔC) of TNK2 sequence were cloned in-frame on the N-terminus of BioID-HA.

2.2 Experimental Design and Statistical Rationale – Three-state stable isotopic labeling with amino acids in cell culture (SILAC) of SW48 parental cells, G12D mutant cells, and G13D mutant cells was performed. Briefly, cells were cultured in RPMI 1640 SILAC media deficient in L-lysine and L-arginine (Thermo Fisher Scientific, Waltham, MA). Parental cells were labeled as "light" by culturing in media supplemented with light lysine (K0) and light arginine (R0). G12D mutant cells were labeled as "medium" by culturing in media supplemented with $^{2}H_4$-Lysine (K4) and $^{13}C_6$-Arginine (R6). G13D mutant cells were labeled as "heavy" by culturing in media supplemented with $^{13}C_6$, $^{15}N_2$-Lysine (K8) and $^{13}C_6$, $^{15}N_4$-Arginine (R10). Three biological replicates were used for performing quantitative analysis of total proteome and phosphotyrosine proteome using high resolution mass spectrometry. Significantly regulated phosphosites and proteins were identified by ANOVA, and a p-value of 0.05 was denoted as the threshold to determine statistically significant changes between two groups.

2.3 Quantitative SILAC proteomics for BioSITe- Two-state stable isotopic labeling with amino acids in cell culture (SILAC) was performed for cells used in this study. Briefly, cells expressing mutant TNK2 ΔC-BirA* were labeled as "light" by culturing in media supplemented with light lysine (K0) and light arginine (R0). Cell expressing TNK2 FL-BirA* were labeled as "heavy" by culturing in media supplemented with 13C6, 15N2-Lysine (K8) and 13C6, 15N4-Arginine (R10). Cells expressing TNK2 FL-BirA* and TNK2 ΔC-BirA* were cultured overnight with 50 μM biotin. Cells were lysed in 8 M urea buffer (20 mM HEPES pH 8.0, 8 M urea, 1 mM sodium orthovanadate, 2.5 mM sodium

pyrophosphate, 1 mM β-glycerophosphate, and 5 mM sodium fluoride), sonicated, and then cleared by centrifugation at 15,000 x g at 4 °C for 20 min. Protein concentration of lysates was determined by BCA Protein Assay. For each biological replicate, equal amounts of protein (10 mg) from each labeling condition was mixed and subjected to trypsin in-solution digestion and BioSITe analysis as described below.

2.4 Immunoprecipitation Experiments and Western blot analysis – Cells were harvested and lysed in modified RIPA buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mm EDTA, 1% Nonidet P-40, 0.25% sodium deoxycholate, and 1 mM sodium orthovanadate in the presence of protease inhibitors and centrifuged. The supernatant was resolved by SDS-PAGE, transferred to nitrocellulose membranes, and probed with primary and horseradish-peroxidase-conjugated secondary antibodies. For validating hyperphosphorylation events, supernatant was subjected to immunoprecipitation with anti-phosphotyrosine antibodies (4G10). After washing, the immunoprecipitates were resolved by SDS-PAGE and assayed by Western blotting using antibodies against proteins of interest.

2.5 siRNA knockdown and proliferation assays – siRNA (10 nM) targeting MPZL1 and KRAS were used for transfections with RNAiMax (Thermo Fisher Scientific, Waltham, MA). Scrambled siRNA sequence was used as a control. Cells were harvested 36 hours post transfection for assessing knockdown efficiency or other follow-up experiments.

2.6 Trypsin Digestion and Peptide Preparation – Following cell culture, peptides were prepared using an in-solution tryptic digestion protocol with slight modifications. In brief, frozen pellets of SILAC-labeled cells were lysed in 8 M urea buffer (20 mM HEPES pH 8.0, 8 M urea, 1 mM sodium orthovanadate, 2.5 mM sodium pyrophosphate, 1 mM β-

glycerophosphate, and 5 mM sodium fluoride), sonicated, and then cleared by centrifugation at 15,000 x g at 4 °C for 20 min. Protein concentration of lysates was determined by BCA Protein Assay. Twenty milligrams of protein lysates from each labeling condition was mixed and subjected to trypsin in-solution digestion. Briefly,the resultant mixture was reduced with 5 mM dithiothreitol and alkylated with 10 mM iodoacetamide. For in-solution tryptic digestion, the resulting protein extracts were diluted in 20 mM HEPES pH 8.0 to a final concentration lower than 2 M urea incubated with 1 mg/mL TPCK-treated trypsin on an orbital shaker at 25 °C overnight. Protein digests were acidified with 1% trifluoroacetic acid (TFA) to quench the digestion reaction and then subjected to centrifugation at 2000 xg at room temperature for 5 min. The resulting supernatants were desalted using SepPak C18 cartridge (Waters Corporation, Milford, MA). Eluted peptides were lyophilized to dryness prior to phosphotyrosine peptide enrichment.

2.7 Affinity Enrichment of Phosphotyrosine Peptides – For phosphotyrosine enrichment, 60 mg per replicate of lyophilized tryptic peptides were used. Phosphotyrosine rabbit monoclonal antibody kit (PTM Scan, P-Tyr-1000; Cell Signaling Technology, Danvers, MA) was used for affinity enrichment of phosphotyrosine peptides. Immunoaffinity purification (IAP) of phosphopeptides was carried out as previously described (50). Briefly, after lyophilization, 60 mg of peptide mixture was dissolved in 1.4 ml of IAP buffer (50 mM MOPS pH 7.2, 10 mM sodium phosphate, 50 mM NaCl) and subjected to centrifugation at 2000 x g at room temperature for 5 min. Before IAP, P-Tyr-1000 beads were washed with IAP buffer twice at 4 °C and the pH of the supernatant containing peptides was adjusted to 7.2 by adding 1 M Tris Base. For IAP, the supernatant

was incubated with P-Tyr-1000 beads at 4 °C for 30 min and the beads were washed three times with IAP buffer and then twice with water. Peptides were eluted twice from beads by incubating the beads with 0.1% TFA at room temperature.

2.8 Basic pH Reversed-phase liquid chromatography – The peptide digest was fractionated using basic pH RPLC for total proteome analysis. Fractionation was carried out at pH 9.5 on a XBridge C18 column (250 × 4.6 mm, 5 µm, 200 A°, Waters Corporation, Milford, MA). Peptides were reconstituted in basic pH reversed-phase liquid chromatography (bRPLC) solvent A (5 mM triethylammonium bicarbonate (TEAB), pH 9.5) and loaded on a C18 column using an Agilent 1260 HPLC system. Peptides were fractionated by a 70 min gradient of 8–50% solvent B (5 mM TEAB, 90% acetonitrile, pH 9.5) at a flow rate of 1 mL/min. The peptide fractions were collected in a 96-well plate. Fractionated peptides were concatenated into 24 fractions, vacuum dried, and stored at −80°C until LC-MS/MS analysis.

2.9 LC-MS/MS analysis– The peptide fractions were loaded on a 2 cm trap column (Acclaim PepMap 100, C18, 5µm particle size, 100 µm i.d. 100 Å pore size, Thermo Scientific, San Jose, CA) using 0.1% formic acid with a flow rate 10 µl/min for 5 minutes. The peptides were separated on a 50 cm analytical column (Acclaim PepMap 100, C18, 3 µm particle size, 75 µm i.d. 100 Å pore size, Thermo Scientific, San Jose, CA) with a 100 min gradient from 7% to 40% acetonitrile in 0.1% formic acid with a flow rate of 300 nl/min. The spray voltage was set to 2.3 kV while capillary temperature was set to 200°C. The samples were analyzed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific, Bremen, Germany). The MS instrument was operated in data-dependent acquisition mode. A survey full scan MS (from m/z 350–1,550) was acquired in the

Orbitrap with a resolution of 120,000 at m/z 200 with a maximum AGC target value of 200,000 ions. The data-dependent MS/MS was carried out using the Top Speed method with a duty cycle of 3 seconds. Singly charged precursor ions were excluded while precursor ions with charge states ≥2 were sequentially isolated to a target value of 50,000 ions and fragmented in the higher-energy collisional dissociation cell using 32% normalized collision energy. The maximum ion injection time for MS and MS/MS were set to 60 ms and 100 ms, respectively. Fragment ion spectra were detected in an Orbitrap mass analyzer with a resolution 30,000 at m/z 200. Dynamic exclusion was enabled with one event of fragmentation followed by the exclusion of the precursor for the next 30 seconds within 10 ppm of the selected m/z. For all measurements with the Orbitrap detector, a lock-mass ion from ambient air (m/z 445.120025) was used for internal calibration (80).

2.10 Mass Spectrometry Data Analysis for pTyr and Global Proteome – Proteome Discoverer software suite (v 2.2; Thermo Fisher Scientific, San Jose, CA) was used for quantitation and database searches. The MS/MS data were searched using the SEQUEST search algorithm against a Human RefSeq database (v73 containing 73,198 entries) supplemented with frequently observed contaminants. Additionally, sequence entries corresponding to mutant KRAS containing G12D or G13D were added to this database. Search parameters used for SEQUEST search algorithm included trypsin as protease with full specific and a maximum of one allowed missed cleavage; carbamidomethylation of cysteine as a fixed modification; oxidation at methionine as a variable modification. For all the 3-State SILAC labeled samples, $^{13}C_6$, $^{15}N_2$-lysine, $^{2}H_4$-lysine, $^{13}C_6$-arginine and $^{13}C_6$, $^{15}N_4$-arginine were considered as variable modifications. For data analysis of

phosphotyrosine enrichments, phosphorylation at tyrosine was additionally selected as a variable modification. The precursor tolerance was set at 10 ppm while the fragment match tolerance was set to 0.05 Da. The PSMs, peptides and proteins were filtered at a 1% false discovery rate cut-off calculated using decoy database searches. For quality control, the probability that an identified phosphorylation was modifying each specific Ser/Thr/Tyr residue on each identified phosphopeptide was determined from the PhosphoRS algorithm (81). Post-processing for phosphotyrosine data was done as follows: After database searching, SILAC ratios of peptides were quantified with PyQuant (81). Resulting SILAC ratios were maintained at the peptide level and grouped by site, and median values were calculated. Additional median normalization was applied to the G12D/Parental ratios and G13D/Parental ratios. Finally, the median values from the three replicates were calculated. A 2-fold cutoff was selected for hyper-phosphorylation, and a 0.5-fold cutoff was selected to denote hypo-phosphorylation.

2.11 BioSITe- Samples were processed using BioSITe method as previously described by Kim et al. (18). Briefly, peptide samples dissolved in BioSITe capture buffer (50 mM Tris, 150 mM NaCl, 0.5% Triton X-100) were incubated with anti-biotin antibody bound to protein-G beads for 2 hours at 4°C. Following incubation, beads were washed multiple times with PBS and then washed two times with BioSITe capture buffer, two times with 50 mM Tris and two times with ultrapure water. Biotinylated peptides were eluted four times using elution buffer (80% acetonitrile and 0.2% trifluoroacetic acid in water). The eluted sample was further cleaned up using C18 reversed-phase column and subject to LC-MS/MS analysis.

2.12 LC-MS/MS analysis for BioSITe – Peptide samples were analyzed on an Orbitrap Fusion Lumos Tribrid Mass spectrometer coupled with the Easy-nLC 1200 nano-flow liquid chromatography system (Thermo Fisher Scientific). Peptides were reconstituted using 20 µL 0.1% formic acid and loaded on an Acclaim PepMap 100 Nano-Trap Column (100 µm x 2 cm, Thermo Fisher Scientific) packed with C18 particles (5 µm) at a flow rate of 4 µl per minute. Peptides were separated using a linear gradient of 7% to 30% solvent B (0.1% formic acid in 95% acetonitrile) at a flow rate of 300-nl/min over 95 min on an EASY-Spray column (50 cm x 75 µm ID, Thermo Fisher Scientific) packed with 2 µm C18 particles, which was fitted with an EASY-Spray ion source that was operated at a voltage of 2.3 kV. Mass spectrometry analysis was performed in a data-dependent manner with a full scan in the mass-to-charge ratio (m/z) range of 300-18,000 in the "Top Speed" setting, three seconds per cycle. MS scans for precursor ion detection were measured at a resolution of 120,000 at an m/z of 200. MS/MS scans for peptide fragmentation ion detection were acquired by fragmenting precursor ions using the higher-energy collisional dissociation (HCD) method and detected at a mass resolution of 30,000, at an m/z of 200. Automatic gain control for MS was set to one million ions and for MS/MS was set to 0.05 million ions. A maximum ion injection time was set to 50 ms for MS and 100 ms for MS/MS. MS was acquired in profile mode and MS/MS was acquired in centroid mode. Higher-energy collisional dissociation was set to 32 for MS/MS. Dynamic exclusion was set to 35 seconds, and singly-charged ions were rejected. Internal calibration was carried out using the lock mass option (m/z 445.1200025) from ambient air.

2.13 Post processing and bioinformatics for BioSITe- Proteome Discoverer (v 2.2; Thermo Scientific) suite was used for identification. Raw files derived from 3 replicate

LC-MS/MS runs, and were searched together. Spectrum selector was used to import spectrum from raw files. During MS/MS preprocessing, the top 10 peaks in each window of 100 m/z were selected for database searches. The tandem mass spectrometry data were then searched using SEQUEST algorithm against protein databases (For BioID experiments; Human RefSeq database (v73 containing 73,198 entries) with the addition of fasta file entries for TNK2 FL-BirA* and TNK2 ΔC-BirA*. The search parameters for the identification of biotinylated peptides were as follows: a) trypsin as a proteolytic enzyme (with up to three missed cleavages); b) minimum peptide length was set to 6 amino acids. c) peptide mass error tolerance of 10 ppm; d) fragment mass error tolerance of 0.02 Da; and e) carbamidomethylation of cysteine (+57.02146 Da) as a fixed modification and f) oxidation of methionine (+15.99492 Da), 13C6, 15N2-lysine (K8),13C6, 15N4-arginine (R10), biotinylation of lysine (+226.077598 Da), and biotinylation of heavy lysine (+234.091797) as variable modifications. Peptides and proteins were filtered at a 1% false-discovery rate (FDR) at the PSM level using Percolator node (82) and at the protein level using protein FDR validator node, respectively.

Identified protein and peptide spectral match (PSM) level data were exported as tabular files from Proteome Discoverer 2.2. We used PyQuant (19) for obtaining the relative quantification of biotinylated peptides with SILAC light or heavy amino acids. To derive the precursor ion abundance values for the isotopic counterparts for each biotinylated peptide identified from our database search, peak area calculated by PyQuant was used. Briefly, PyQuant scans for the light and heavy peptide (K8R10) peaks in the full MS1 scan and calculates the peak area over the extracted ion chromatogram (XIC). Precursor ion abundances thus computed were used for relative abundance estimates of

biotinylated peptides. Missing abundance values in any LC-MS/MS experiment were replaced with the minimum value. We used an in-house Python script to compile the peptide level site information mapped to UniProt or RefSeq sequences. The summary count on the number of biotinylated sites, supported peptides and PSMs and the relative fold-change information were then calculated at the peptide and protein levels.

2.14 Data Availability – All mass spectrometry datasets acquired for the study described in Chapter 3 was deposited to ProteomeXchange (http://proteomecentral.proteomexchange.org) and is available via the PRIDE database with the accession number PXD009843. Similarly, the data corresponding to the study described in Chapter 4 is available via the PRIDE database with the accession number PXD010474.

CHAPTER 3.

MUTATION-SPECIFIC AND COMMON PHOSPHOTYROSINE SIGNATURES OF *KRAS* G12D AND G13D ALLELES

3.1 KRAS G12D or G13D mutant cells exhibit distinct features

Phosphotyrosine signaling has a well-established role in sustaining oncogenic signaling, as reflected by the number of targeted therapies directed against molecules in tyrosine kinase signaling pathways (83). However, no controlled studies have yet defined the phosphotyrosine landscape associated with individual KRAS mutations. As a part of an effort to systematically explore the signaling differences between two frequently occurring KRAS mutations, G12D and G13D **(Figure 3.1A)**, we examined their tyrosine phosphorylation in an isogenic system using parental SW48 human colorectal cells (KRAS wild-type) and variants of these cells engineered to harbor two common KRAS mutations – G12D or G13D (6).

The isogenic pair of KRAS mutant cells harboring either G12D or G13D mutations exhibit distinct cell morphology **(Figure 3.1B)**. While G12D mutant cells form cell clusters with a cobblestone-like morphology, G13D mutants exhibit a fibroblastic morphology with cell projections. A previous study has reported that SW48 cells with the G12D mutation show increased resistance to EGFR inhibitors than cells expressing the G13D mutation (58). To determine whether our isogenic system also modeled this variation in sensitivity to EGFR inhibition, we performed cell proliferation assays. In agreement with previous studies, we found that KRAS G12D mutant cells were more resistant to erlotinib than parental or G13D mutant cells **(Figure 3.1C)**.

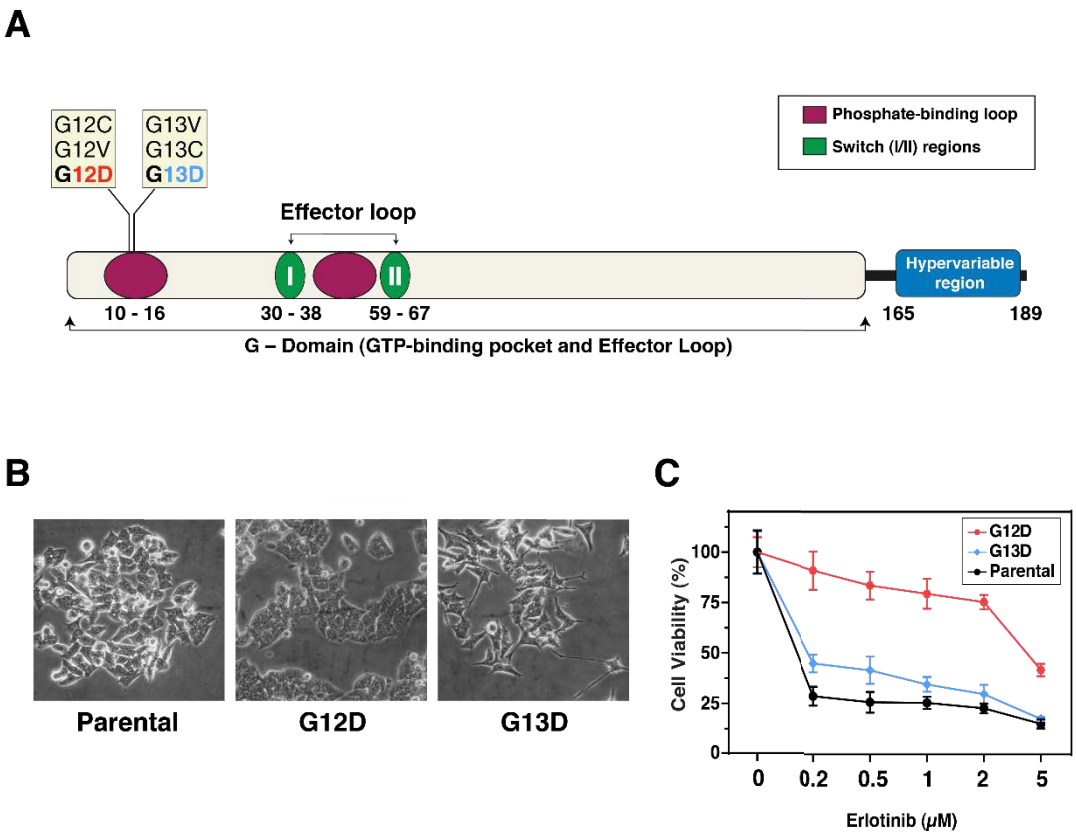

Figure 3.1 Distinct phenotypes observed in codon 12 mutations in KRAS. (A) Schematic depicting the domain architecture of KRAS protein along with the location of two mutational hotspots. Various domains are indicated as colored ovals and the hypervariable region is shown in blue at the C-terminus. Amino acid positions of the domains are marked. (B) A photomicrograph of parental SW48 colorectal cancer cells or those engineered to contain heterozygous G12D or G13D mutations. (C) Effect of treatment with the EGF receptor inhibitor, erlotinib. The parental and mutant (G12D or G13D) cells are shown with different colors as indicated.

3.2 Tyrosine phosphorylation and total proteome landscape in isogenic G12D and G13D cells

To determine whether G12D and G13D KRAS mutations induce different tyrosine phosphorylation changes, we performed an initial survey of the global phosphotyrosine profile of each mutant using Western blot analysis with an anti-phosphotyrosine antibody. We observed that the phosphotyrosine profiles of the two mutants showed distinct phosphotyrosine bands with some shared features **(Figure 3.2A)**. In addition, relative to parental cells, those carrying a G12D mutation exhibited substantially higher tyrosine phosphorylation levels than those harboring the G13D mutation. To identify proteins that show altered abundance and phosphorylation, we employed a quantitative proteomics approach using stable isotope labeling with amino acids in cell culture (SILAC) (84) by growing our cell lines in media supplemented with light or stable heavy isotope labeled L-lysine and L-arginine amino acids. After culturing cells in respective media, WT (K0, R0), G12D (K4, R6), and G13D (K8, R10) cells were combined, lysed and digested with trypsin. For total proteome analysis, peptides were pre-fractionated into 24 fractions. For phosphotyrosine analysis, phosphotyrosine peptides were enriched using anti-phosphotyrosine antibodies. These experiments were carried out using biological triplicates. LC-MS/MS analysis was carried out on an Orbitrap Fusion Lumos Tribrid mass spectrometer **(Figure 3.2B)**. Total proteins and phosphopeptides were identified using a 1% false discovery rate cutoff. Phosphosites were considered localized if a tyrosine residue was assigned with ≥75% probability as calculated by the PhosphoRS algorithm (81). Our analysis of the total proteome resulted in the identification of 8,463 proteins. These proteins included 166 kinases, 82 phosphatases and 318 transcription factors. Total proteome

analysis confirmed the expression of the expected KRAS mutations as we detected MS/MS spectra for the isotopically-labeled mutant peptides derived from G12D **(Figure 3.2C)** or G13D KRAS proteins **(Figure 3.2D)**. The detection of peptides corresponding to WT KRAS allele in cell expressing G12D and G13D further confirmed the heterozygous nature of the mutations in the cells.

The phosphotyrosine experiments led to the collective identification of 2,052 unique phosphotyrosine sites (derived from 667 proteins). Quantitation was performed by the program PyQuant (85) which provided quantitation of 875 phosphotyrosine sites (derived from 555 proteins), of which 659 sites (derived from 440 proteins) were quantified in all three replicates. MS/MS spectra for select phosphotyrosine-modified peptides were manually examined to assess the quality of identification. The phosphotyrosine data included 80 phosphotyrosine sites derived from 51 kinases, 74 sites in 39 signaling adaptor/scaffold proteins and 22 sites in 12 transcription factors. In agreement with the Western blot analysis, the distribution pattern for phosphopeptide fold-changes for each KRAS mutant relative to WT suggested that the extent of tyrosine phosphorylation was indeed higher in G12D cells than G13D or parental cells **(Figure 3.2E)**. Relative to the parental cells, G12D-harboring cells exhibited hyperphosphorylation at 53 phosphotyrosine sites (corresponding to 44 proteins), and G13D cells showed an upregulation of 60 (corresponding to 57 proteins). Overall, our data demonstrate that despite the previously reported characterization of the global phosphoproteome in the same model system (25), our focused analysis of the tyrosine phosphorylated fraction of the phosphoproteome revealed many mutant KRAS-induced signaling events that were previously uncharacterized. We examined our data for sites that were hyperphosphorylated

in one of the mutants but were unchanged or downregulated in the other. Of the 199 phosphotyrosine sites within our data that showed a KRAS mutation-specific pattern in their hyperphosphorylation, 110 phosphotyrosine sites (corresponding to 88 proteins) show increased phosphorylation only in G12D, and 89 sites (corresponding to 83 proteins) upregulated only in G13D.

3.3 Mutant *KRAS*-induced changes in the tyrosine phosphoproteome

Our study sought to determine what tyrosine phosphorylation events were differentially regulated downstream of KRAS mutations in general. To do this, we examined the phosphotyrosine sites showing significant changes (>2 folds and $P<0.05$) in phosphorylation in both KRAS mutants **(Figure 3.3A – B)**. Of the identified differentially tyrosine phosphorylated peptides induced by KRAS mutations, 12 phosphotyrosine sites (derived from 11 proteins) displayed hyperphosphorylation in both G12D and G13D mutant cells **(Figure 3.3C)**. Of the sites with increased tyrosine phosphorylation in both KRAS mutants as compared to the parental cells, a subset are derived from kinases, including DYRK2 (Y382), FYN (Y185), LYN (Y364), and TNK2 (Y827). Non-receptor tyrosine kinase 2 (TNK2) is a cytosolic effector of activated transmembrane RTKs and is involved in promoting cancer progression (86–88), were hyperphosphorylated in both KRAS mutants. Phosphorylation of Y827, which is located in the EGFR binding domain of TNK2, is known to be mediated by Src kinases (89). Representative spectrum data for this phosphosite is shown in **Figure 3.3D**. Furthermore, immunoprecipitation using phosphotyrosine specific antibodies and subsequent Western blot analysis for TNK2 confirmed that overall tyrosine phosphorylation of TNK2 was increased in both KRAS mutants **(Figure 3.3E)**. In addition to kinases, phosphosites in several calcium-binding

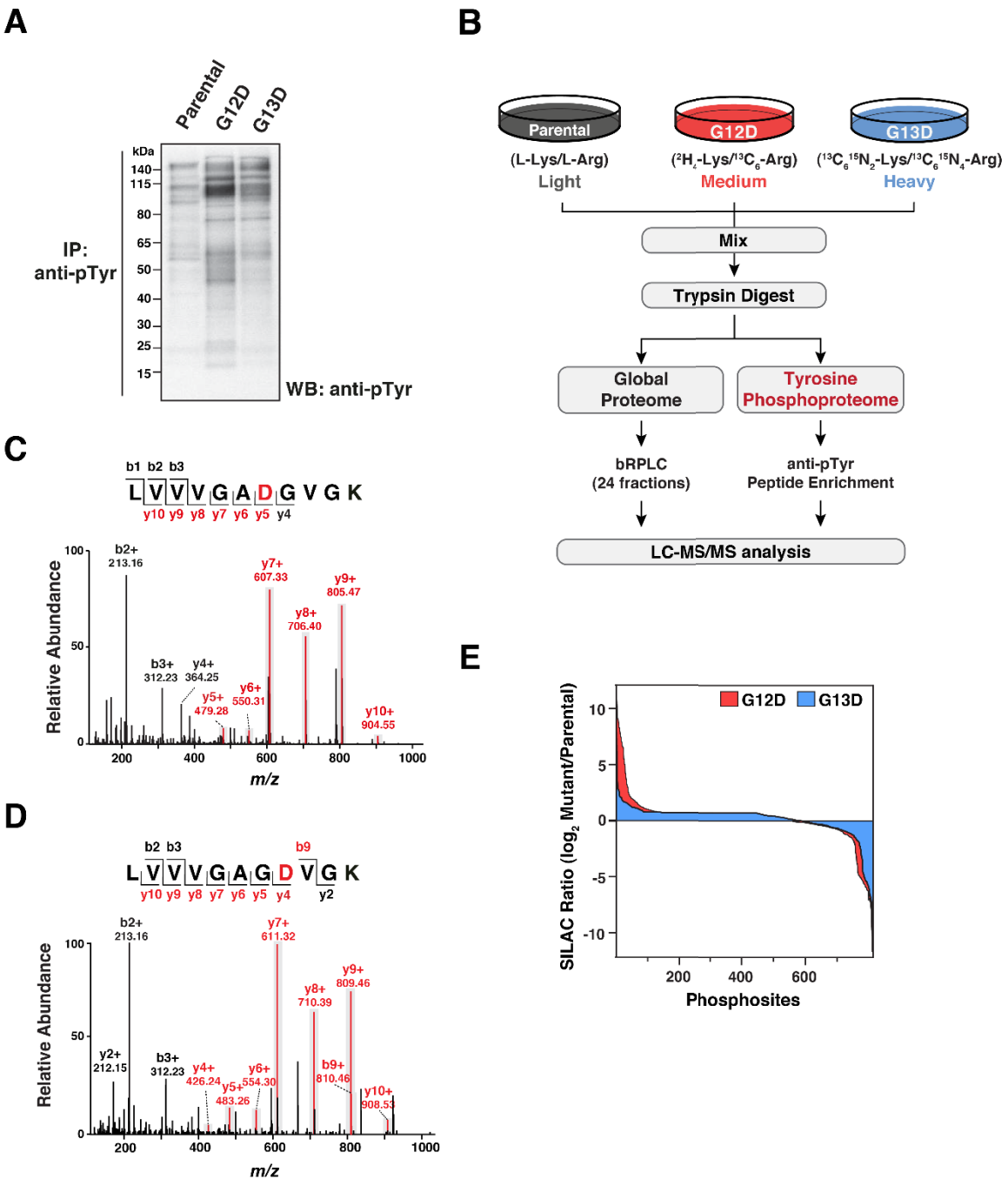

Figure 3.2 Global and phosphotyrosine profiling of KRAS mutant cells. (A) Global phosphotyrosine profile of parental and mutant (G12D or G13D) cells revealed by Western blot analysis. (B) Schematic of the experimental strategy used for quantitative profiling of phosphotyrosine and total proteomes. (C-D) Representative MS/MS spectra showing identification of peptides corresponding to G12D mutant (C) and G13D mutant (D) as indicated. Fragment ions confirming the presence of amino acid substitutions are indicated in red. (E) Distribution of log2-transformed ratios for tyrosine phosphopeptides (G12D/Parental and G13D/Parental).

proteins, including ANXA2 (Y48, Y256), CALM2 (Y100), and ITSN2 (Y968), were hyperphosphorylated in both KRAS mutants. Calmodulin (CALM2) has a well-established role in mutant KRAS signaling and is also a known RAS binding partner (13,90,91). Further studies need to be conducted to investigate the particular signaling events leading to these tyrosine hyperphosphorylation events downstream of mutant KRAS. We hypothesize that these changes downstream of mutant KRAS proteins may, in part, be a result of differential binding of canonical wild-type KRAS effectors, or binding and recruitment of novel downstream effectors.

3.4 Mutant KRAS-induced changes in the total proteome

In addition to phosphotyrosine signaling, we also sought to comprehensively map the changes in protein expression in the presence of mutant KRAS. Our quantitative analysis revealed that relative to parental cells, 134 and 182 proteins were significantly more abundant (2-fold or higher; $p \leq 0.05$) in G12D **(Figure 3.4A)** and G13D **(Figure 3.4B)** cells, respectively. Of these, 44 proteins were upregulated downstream of both KRAS mutations **(Figure 3.4C)**. This set of proteins included several cell surface molecules such as surface proteoglycan/glycoproteins (CD99 and GPC1), tetraspanins (CD63 and CD151) and TNF receptor superfamily member 12A (TNFRSF12A). This included upregulation of CD73, a GPI-anchored enzyme, which catalyzes the conversion of AMP to adenosine and is being pursued as a target for cancer immunotherapy. Upregulation of CD73 in cells/tumors expressing mutant KRAS has been previously reported by multiple studies including proteomic studies that reported CD73 upregulation in MCF10A cells expressing the KRAS G12V mutant (92). Overexpression of this protein in human colorectal cancer tumors is correlated with a poorer survival (93,94). Other upregulated molecules included

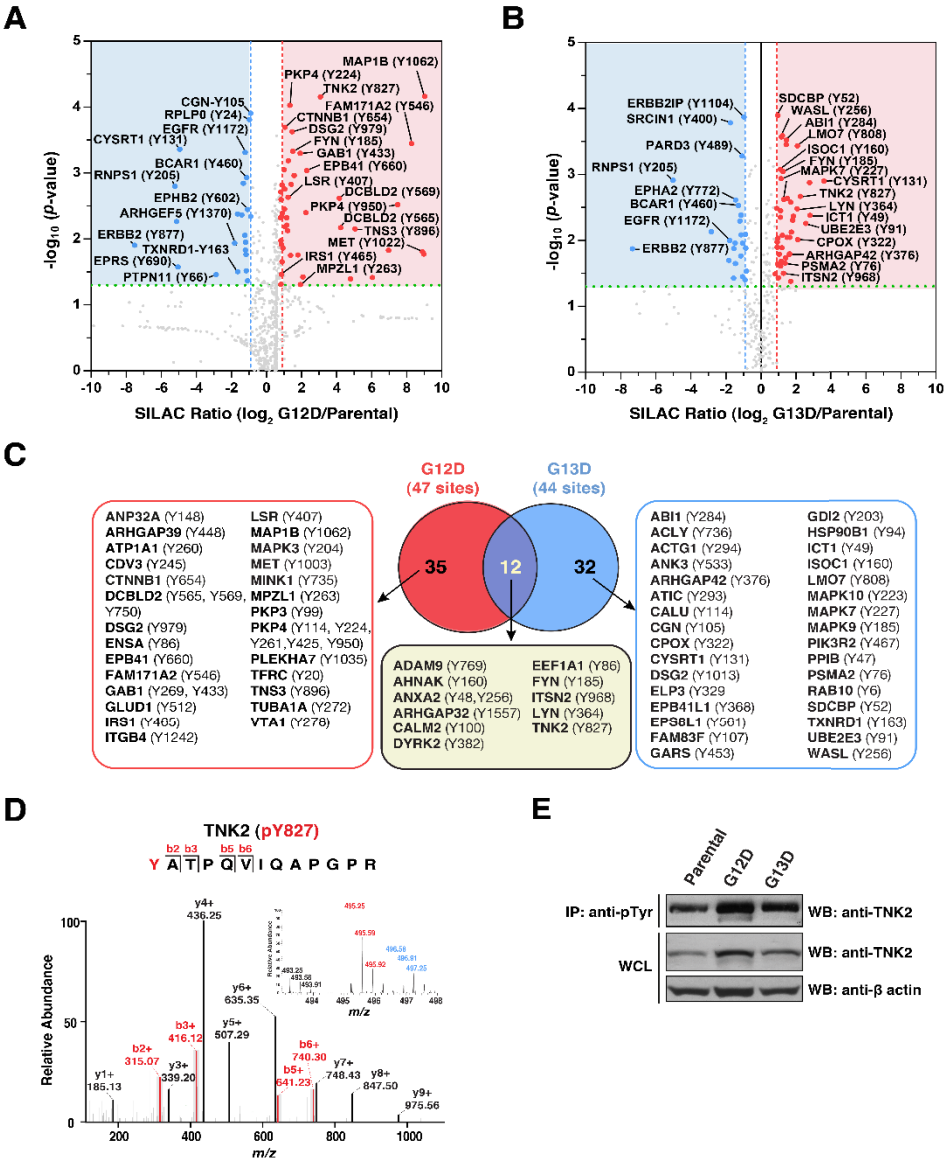

Figure 3.3. Mutant KRAS-induced changes in the tyrosine phosphoproteome. (A-B) Volcano plots depicting relative differences in phosphotyrosine sites between parental cells and (A) G12D mutated and (B) G13D mutated cells. The ratio of relative intensity of phosphotyrosine peptides in individual KRAS mutants versus parental cells is shown on the x-axis (log2-transformed). The y-axis shows the p-values (log2-transformed). The vertical red/blue lines indicate 2-fold changes in the abundance ratios and the horizontal green line signifies a p-value cutoff of 0.05. Phosphosites showing statistically significant (p-value ≤ 0.05) hyper- or hypo-phosphorylation are marked in red or blue, respectively. Selected phosphoproteins/sites are labeled. (C) Venn diagram showing the set of tyrosine phosphorylation sites hyperphosphorylated in both or individual KRAS mutant cells (D) Representative MS/MS spectrum of phosphopeptide corresponding to TNK2 pY827. Fragment ions confirming the localization of the phosphotyrosine site are indicated in red. Representative MS1 level spectrum used for quantifying relative abundance in parental, G12D mutant (red), and G13D mutant (blue) cells is shown as an inset. (E) Western blot analysis of anti-phosphotyrosine immunoprecipitates from parental and mutant cells using protein-specific antibodies against non-receptor tyrosine kinase 2 (TNK2).

proteins associated with DNA replication, recombination, and repair, specifically members of the mini chromosome maintenance complex (MCM complex) which has a role in DNA unwinding (MCM2, MCM3, MCM4, MCM5 and MCM6). Both KRAS mutants showed upregulation of DNA damage response activators, including Fanconi anemia complementation group G (FANCG), MHC class I polypeptide-related sequence A (MICA), and macrophage migration inhibitory factor (MIF). Consistent with the known role of mutant KRAS in driving metabolic reprogramming of cancer cells (95,96), we also observed upregulation of enzymes involved in fatty acid metabolism (UBE2L6, GPD2 and ACSL5). Finally, several upregulated proteins identified in our analysis have not been previously reported or characterized in the context of mutant RAS-induced signaling, such as transmembrane protein 2 (TMEM2), GTP cyclohydrolase I feedback regulatory protein (GCHFR), Kallikrein Related Peptidase 10 (KLK10), FAM122B, KIAA1524, and gasdermin D (GSDMD). We used Western blot analysis to validate some of the common protein alterations observed in our MS data including CD73 (5'-nucleotidase ecto), HSPB1 (heat shock protein beta-1) and KRAP (KRAS-induced actin-interacting protein) **(Figure 3.4D)**. Overall, our global proteome analysis of isogenic KRAS mutant cells represents a resource for understanding the protein expression changes that are induced upon endogenous expression of two of the most common KRAS mutations in cancer.

3.5 Unique tyrosine hyperphosphorylation events downstream of KRAS G12D and G13D

We observed 110 phosphotyrosine sites in 88 proteins to be hyperphosphorylated in G12D, that were unchanged or decreased in G13D **(Figure 3.5A)**. This signature

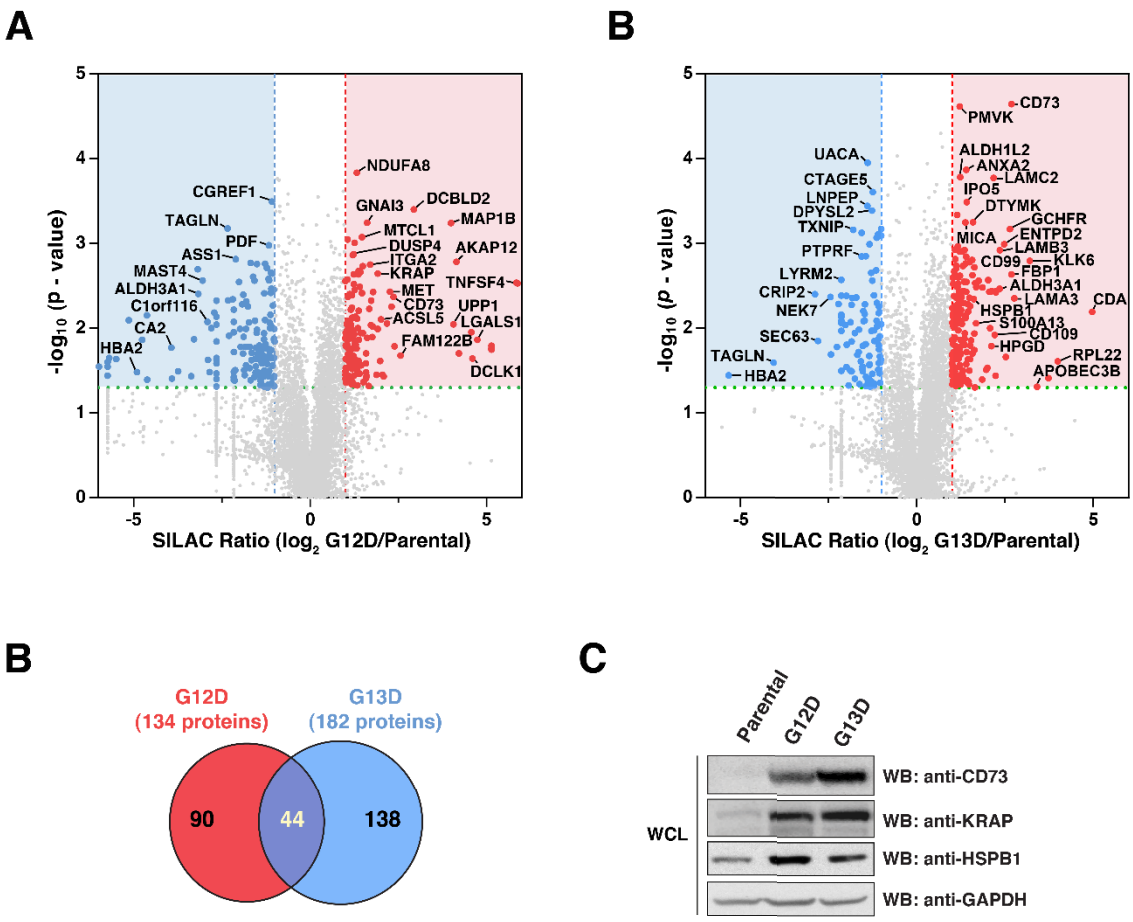

Figure 3.4 – Mutant KRAS-induced signatures in the total proteome. (A-B) Volcano plots depicting relative differences in total protein abundances between parental cells and (A) G12D mutated and (B) G13D mutated cells. The ratio of relative intensity of a protein in individual KRAS mutants versus parental cells is shown on the x-axis (log2-transformed). The y-axis shows the p-values (log2-transformed). The vertical red/blue lines indicate 2-fold changes in the abundance ratios and the horizontal green line signifies a p-value cutoff of 0.05. Proteins showing statistically significant (p-value $\leq$ 0.05) up-regulation or down-regulation are marked in red or blue, respectively. Selected proteins are labeled. (C) Venn diagram showing the set of proteins showing increased abundance in both or individual KRAS mutant cells relative to parental cells. (D) Western blot analysis of whole cell lysates from parental and mutant cells using a protein-specific antibodies directed against 5'-nucleotidase ecto (CD73), KRAS-induced actin-interacting protein (KRAP), and heat shock protein beta-1 (HSPB1).

included hyperphosphorylation of sites in kinases, including MET (Y1003), MAPK3 (Y204), and MINK (Y735). Several G12D-induced phosphorylation changes involved proteins and/or phosphosites that are not well-studied. For instance, hyperphosphorylation of Y546 in FAM171A2 (family with sequence similarity 171, member A2), a poorly characterized integral membrane protein, was observed only in G12D mutant cells. We also observed hyperphosphorylation of discoidin, CUB, and LCCL domain containing 2 (DCBLD2) at Y715, a site that was reported to undergo phosphorylation mediated by upstream Src tyrosine kinases, resulting in a binding site for the SH2 domain of CRKL (97). One of the G12D-specific features was the hyperphosphorylation of myelin protein zero-like 1 (MPZL1) protein (Y263). The MPZL1 protein is a single-pass transmembrane glycoprotein involved in extracellular matrix-induced signaling and regulation of cell adhesion. Increased phosphorylation was observed at Y263, a site localized within the cytoplasmic Immunoreceptor Tyrosine-based Inhibitory Motif (ITIM) that is known to recruit and activate protein tyrosine phosphatase, non-receptor type 11 (PTPN11) (98–101). Phosphorylation of this site is regulated by c-Src and controls the activation status of MPZL1 (102). To validate the G12D-specific hyperphosphorylation of this site as indicated by our quantitative MS data, Western analysis using an antibody specific for phosphorylated Y263 was performed. As shown in **(Figure 3.5C)**, our analysis revealed hyperphosphorylation of MPZL1 at Y263 specifically in G12D cells but not in G13D cells. In addition, multiple proteins involved in focal adhesion and adherens junctions exhibited G12D-specific hyperphosphorylation at one or more sites. These included CTNNB1 (Y654), PKP3 (Y99), PKP4 (Y114, Y224, Y261, Y425, Y950), PLEKHA7 (Y1035), and TNS3 (Y896). Finally, our data also highlighted other notable proteins showing

hyperphosphorylation in this mutant including enhancer of rudimentary homolog (ERH-Y92) a protein shown to be synthetically lethal in mutant KRAS cells (45). Collectively, our findings from the analyses of the tyrosine phosphoproteome bring to attention several G12D mutation-specific phosphorylation events that warrant further investigation for their potential involvement in G12D-specific oncogenic phenotype.

To understand the tyrosine phosphorylation changes occurring uniquely downstream of the G13D mutation, we examined the 32 phosphotyrosine sites (corresponding to 32 proteins) that were hyperphosphorylated only in G13D cells. Among these sites were those mapping to MAPK family kinases, which included MAPK7 (Y227), MAPK9 (Y185), and MAPK10 (Y223). Pathway analysis revealed G13D-specific hyperphosphorylation of many proteins involved in catalyzing diverse metabolic cellular processes. This group included phosphosites in ACLY (Y736), ATIC (Y293), CALU (Y114), CPOX (Y332), ISOC1 (Y106), TXNRD1 (Y163), and UBE2E3 (Y91). Other tyrosine hyperphosphorylation events induced by the G13D mutation involved proteins in regulating GTP/GDP binding (ARHGAP42 (Y376), GDI2 (Y203), RAB10 (Y6), WASL (Y256)), RNA translation (GARS (Y453), ICT1 (Y49)), and protein processing and folding (HSP90B1 (Y94), PPIB (Y47), PSMA2 (Y76)). Overall, our analysis reveals that the phosphotyrosine changes induced by the G13D mutation differed significantly from those induced downstream of G12D mutation, and mostly impacts proteins involved in the regulation of enzymatic activity and assorted metabolic process.

3.6 Unique protein expression changes downstream of KRAS G12D and G13D

To identify proteins that exhibited significant changes in abundance downstream of each KRAS mutation, we focused on proteins that showed a ≥2-fold change ($P<0.05$) between the mutants ((G12D/Parental)/(G13D/Parental)) **(Figure 3.5B)**. Within this set of

proteins (n = 356), we found 47 proteins that showed statistically significant ($p < 0.05$) upregulation only in G12D, which included several cell surface proteins such as CD44, CD59, EPCAM, ITGA2, ITGA3, ITGB1, LGALS1 and MET. Western blot analysis confirmed the upregulation of Hepatocyte growth factor receptor (MET), galectin-1 (LGALS1), and Heparan sulfate proteoglycan (CD44) **(Figure 3.5D)**. In addition to MET receptor tyrosine kinase, we found several other kinases that showed G12D-specific upregulation, including calcium/calmodulin dependent protein kinase II delta (CAMK2D), protein kinase N1 (PKN1), Rho associated coiled-coil containing protein kinase 2 (ROCK2), serine/threonine kinase 4 (STK4) and TANK binding kinase 1 (TBK1). Interestingly, TBK1 is a known synthetic lethal partner of oncogenic KRAS, which has been demonstrated to be essential for survival of cells harboring mutant KRAS (103).We also observed G12D-specific upregulation of proteins that are part of multi-protein complexes, such as subunits of heterotrimeric G-protein complex (GNAI1, GNAI2, GNAI3, GNAS), components of the NDC80 kinetochore complex (NDC80, SPC24, SPC25), trafficking protein particle complex (TRAPPC2L, TRAPPC4, TRAPPC5), TTT complex (TTI1, TTI2) and cavins (CAVIN1, CAVIN3). However, the large majority of the proteins showing G12D-specific upregulation were involved in metabolic processes and pathways, including those that drive glycolysis (ENO2, PFKP, GPC4, G6PD, GNE,

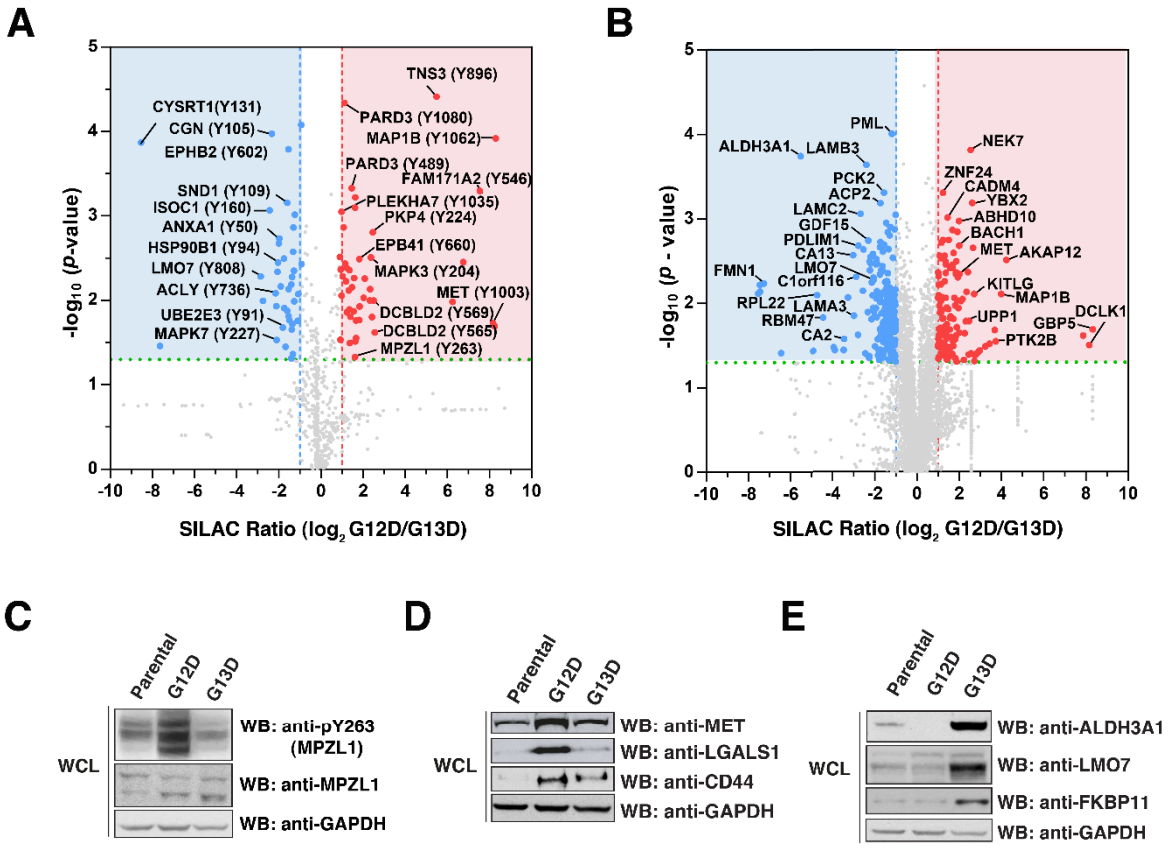

Figure 3.5 KRAS point mutation-specific rewiring of phosphotyrosine proteome in isogenic cells. (A) Volcano plot depicting relative differences in phosphotyrosine peptides in the two KRAS mutant cells. The ratio of relative intensity of phosphotyrosine peptides in G12D versus parental to the relative intensity of phosphotyrosine peptides in G13D versus parental is shown on the x-axis (log2-transformed). The y-axis shows the p-values (log2-transformed). The vertical red line indicates 2-fold changes in the ratios and the horizontal green line signifies a p-value cutoff of 0.05. Phosphosites showing statistically significant (p-value ≤ 0.05) hyper- or hypo-phosphorylation are marked in red or blue, respectively (B) Volcano plot depicting relative differences in total protein abundances in the two KRAS mutant cells. The ratio of relative intensity of a protein in G12D versus parental to the relative intensity of a protein in G13D versus parental is shown on the x-axis (log2-transformed). (C) Western blot analysis of whole cell lysates from parental and mutant cells using a phosphosite-specific antibody directed against Y263 of MPZL1. (D) Western blot analysis of whole cell lysates from parental and mutant cells using a protein-specific antibody directed against Hepatocyte growth factor receptor (MET), galectin-1 (LGALS1), and Heparan sulfate proteoglycan (CD44). (E) Western blot analysis of whole cell lysates from parental and mutant cells using a protein-specific antibody directed against Aldehyde Dehydrogenase 3 Family Member A1 (ALDH3A1), LIM domain only protein 7 (LMO7), and FKBP family of peptidyl-prolyl cis/trans isomerases (FKBP11).

ME2) or are involved in fatty acid/lipid metabolism (ABCD3, ABHD10, ACAD9, ACSL3, CRAT, ECHDC1, ENO2, IDH2, ME2, MECR, PGD, PTGR1 and ZADH2). Further, upregulation of proteins associated with the inner mitochondrial membrane (BPHL, CNP, COX20, MRPS16, PISD, TIMM23, TOMM34 and UQCRQ), and other proteins with general oxidoreductase activity (ACADSB, ALDH4A1, ALDH6A1, BLVRA, FOXRED2, G6PD, MICAL1, NDUFA13, NDUFA8, NDUFB9, PTGR1, SCD and VKORC1) were unique to cells harboring the G12D mutation. Finally, G12D-specific upregulation of several transcription factors (CSRP2, HDAC1, MBD2, PMF1, TAF5, TRIM26, TSC22D1) was also observed. Analysis of proteins showing a KRAS G13D mutation-specific increase in protein abundance (n = 68) included cell surface markers (CD109, FAS, SDC1), cytokines (FAM3C, IL18), several kinases (CAMK1, MAP3K6, ROCK1, RPS6KA3) and transcription factors (FHL2, LMO7, PML, SMAD3, STAT1, ZNF185). Components of the laminin complex (LAMA3, LAMB3, LAMC2) were also among the proteins preferentially upregulated in G13D mutated cells. Western blot analysis confirmed the upregulation of aldehyde dehydrogenase 3 family member A1 (ALDH3A1), galectin-3 (LGALS3), and FK506 binding protein 11 (FKBP11) in G13D cells **(Figure 3.5E)**. G13D-specific upregulation of ALDH3A1 was also observed previously and served as a positive control (77). Overall, our analysis provides an in-depth characterization of differences in the protein expression changes induced by each KRAS mutation, and highlights molecules and pathways that may be of specific importance in sustaining oncogenic signaling in a mutation-specific manner.

3.7 MPZL1, an ill-characterized cell surface glycoprotein, is a novel molecule downstream of mutant KRAS

Some of the tyrosine phosphorylation events that are uniquely induced by either KRAS mutation may be crucial for maintaining their oncogenic signaling. It is also possible that proteins showing increased tyrosine phosphorylation downstream of the G12D mutation may be involved in conferring resistance to EGFR inhibitors. To test this, we performed siRNA knockdown of MPZL1, a relatively understudied cell surface glycoprotein that showed G12D-specific tyrosine hyperphosphorylation in our analysis **(Figure 3.5C)**. Analysis of publicly available datasets for genes that show synthetic lethality (37,40) with mutant KRAS revealed that compared to wild-type cells, KRAS mutated cells are significantly more dependent on expression of MPZL1 **(Figure 3.6A)**. To assess the importance of MPZL1 in parental or mutant cells, we performed a cell viability assay using cells transfected with either control/scramble siRNA or siRNA targeting MPZL1. Knockdown of MPZL1 led to decreased viability of G12D mutant cells specifically, suggesting a selective importance of this protein in sustaining G12D mutant cells **(Figure 3.6B)**. As a positive control, we assessed changes in viability of parental and mutant cells transfected with control or KRAS targeting siRNA. As expected with oncogene addiction/dependency (104), knockdown of KRAS resulted in decreased viability of G12D and G13D mutant cells, but not parental cells **(Figure 3.6C)**. Finally, we assessed the impact of MPZL1 knockdown on each cell line's response to the EGFR inhibitor erlotinib by treating cells transfected with either control or MPZL1 siRNA with varying concentrations of erlotinib. Relative to control siRNA transfected cells, we observed a

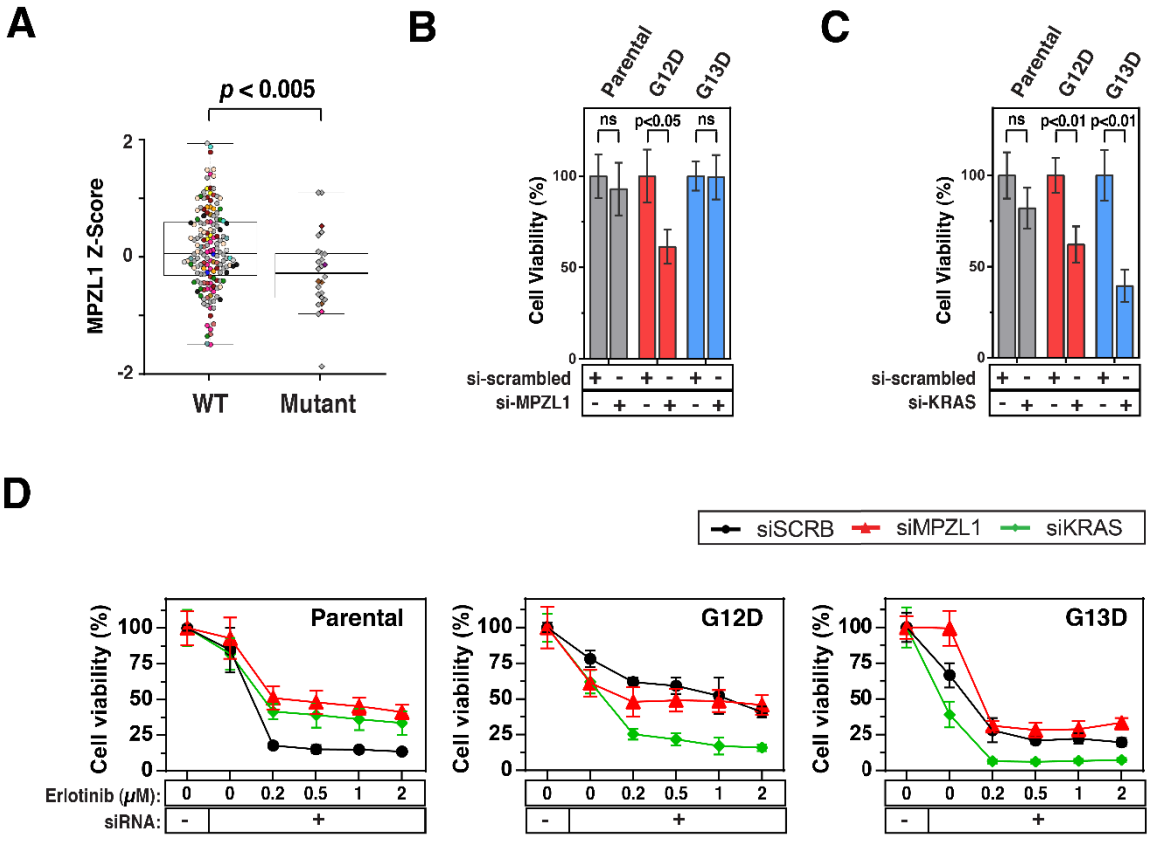

Figure 3.6. Knockdown of MPZL1 leads to decreased cell viability and increased sensitivity in G12D mutant cells. (A) Cells harboring KRAS mutations are significantly more sensitive to MPZL1 knockdown compared to wild-type cells. Data and plot were obtained from genetic dependency datasets available through the CancerGD resource (www.cancergd.org) (40). (B) Cell proliferation assays using parental and mutant cells transfected with either control siRNA (si-scrambled) or siRNA against MPZL1 (si-MPZL1). Percentage cell viability was calculated relative to cells treated with control siRNA. Results for each group were derived from four replicates. P-value was calculated using a paired t-test. (C) Cell proliferation assays using parental and mutant cells transfected with either control siRNA (si-scrambled) or siRNA against KRAS (si-KRAS). (D) Cell proliferation assays using erlotinib-treated parental and mutant cells that were transfected with either control siRNA (si-scrambled), siRNA against MPZL1 or KRAS.

targeted decrease in viability of erlotinib-treated G12D mutant cells that were transfected with siRNA against MPZL1 suggesting an increase in sensitivity of G12D mutant cells to EGFR inhibition **(Figure 3.6D)**. Collectively, these preliminary studies suggest a targeted functional importance of MPZL1 in signaling downstream of the KRAS G12D mutation **(Figure 3.7)**. Although further studies are needed to describe and assess the role of MPZL1 in normal and mutant RAS signaling, our analysis underscores this molecule as a potential therapeutic candidate in KRAS, specifically G12D, mutated cancers.

3.8 Conclusions

Although KRAS, one of the earliest discovered proto-oncogenes, has been the focus of a large number of studies over the last few decades, several aspects of its biology including the subtle biological differences present between the point mutations, still remain a mystery. In this study, we chose an approach that not only permitted us to specifically determine the signaling differences between two common mutations, G12D and G13D, but also allowed a deeper understanding of signaling events impacted by mutant KRAS in general. Enrichment of phosphotyrosine-containing peptides allowed us to probe phosphotyrosine signaling pathways, which have become important therapeutic targets in cancer. Incorporation of a SILAC strategy allowed us to more precisely characterize the subtle differences in tyrosine phosphorylation-mediated signaling downstream of the two KRAS mutations. Finally, the power of this unbiased approach was demonstrated in our study through the identification of numerous unsuspected and novel molecules that, owing to their low abundance or limitations of previous approaches, have not been linked with oncogenic RAS signaling. With the advent of CRISPR/Cas9 technology, the ability to quickly develop isogenic model systems engineered to harbor various mutations/fusions of

interest has become widely accessible. We believe that the coupling of our approach to genomic engineering through CRISPR/Cas9 technology is an excellent strategy for investigating signaling by related oncogenes.

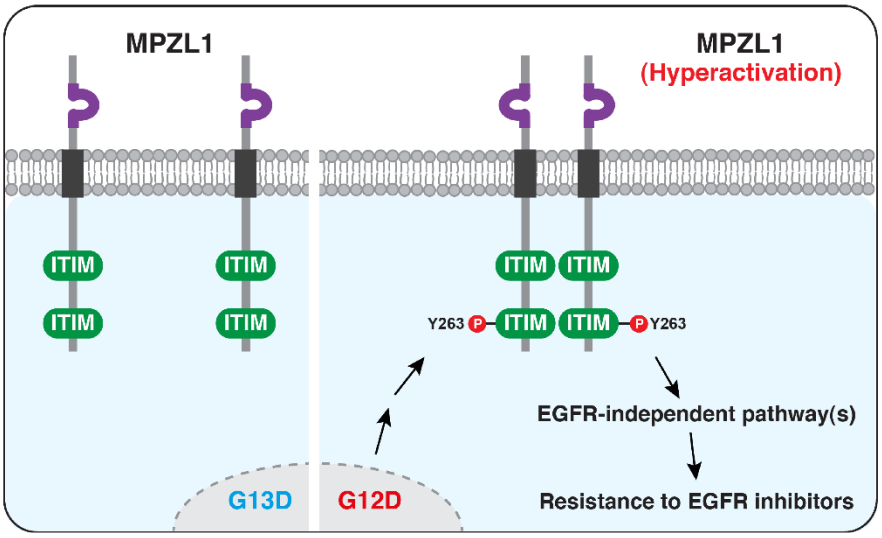

Figure 3.7 Schematic model depicting MPZL1 as a signaling node in KRAS G12D mutation-specific signaling. Select protein domains of MPZL1, such as extracellular Ig-like domain (purple), transmembrane domain (grey) and immunoreceptor tyrosine-based inhibition motif (ITIM), are shown.

CHAPTER 4.

PROXIMITY-DEPENDENT BIOTINYLATION TO ELUCIDATE THE INTERACTOME OF TNK2 NON-RECEPTOR TYROSINE KINASE

The dynamic assembly and disassembly of protein complexes is a fundamental principal that drives biological processes. Compared to other non-receptor tyrosine kinases, reported studies on TNK2 are specifically lacking in focus on characterizing its intracellular molecular binding partners and their contribution to the role of TNK2 in normal and oncogenic signaling. Herein, we describe how we performed a quantitative analysis of protein biotinylation with a site-level resolution as a means to identify TNK2 proximal proteins in the context of a breast cancer cell line.

4.1 Application of BioSITe for the discovery of TNK2 interactors

We took advantage of the BioID system to identify interactors of TNK2. For this, we generated a full-length TNK2 construct fused with BirA* at C-terminal of TNK2 **(Figure 4.1A)**. We also generated a TNK2 C-terminal deletion BirA* fusion construct with the deletion of the C terminal moiety of TNk2 (ΔC) that was lacking important regions that mediate TNK2 interaction with other proteins, including the Proline-rich domain, the EGFR-binding domain (EBD), and the ubiquitin association domain (UBA) in the C-terminus of TNK2. This construct also served as a control to allow us to more clearly differentiate the interacting partners of the functional full-length TNK2 protein. The full length TNK2 (FL) and truncated TNK2 (ΔC) constructs were respectively transfected into HEK293T cells for packaging viral particles that were used to infect HCC1395 cells and stable cell lines were generated by puromycin selection. Relative expression levels of BirA*-tagged TNK2 FL, and ΔC were analyzed by immuno-blot with both anti-HA tag and anti-TNK2 antibodies, showing the different molecular weights of TNK2 protein variants. Western blot analysis using an HA tag and TNK2 specific antibody confirmed the expression of both constructs at their expected molecular weights **(Figure 4.1B)**. To

confirm an increase in global intracellular biotinylation in cells stably expressing TNK2 BirA* fusion constructs, we performed western blot analysis using anti-biotin antibody on whole cell lysates from stable cells cultured in growth medium with or without exogenous biotin overnight. As expected, relative to parental cells and cells not treated with biotin, our BirA* cells incubated with biotin showed an increased global biotinylation signature, which included visibly increased biotinylation of the TNK2-BirA* fusion proteins themselves **(Figure 4.1C)**. To systematically map the molecular interactors specific to the canonical full-length TNK2 protein, we opted for a quantitative approach by incorporating stable isotope labeling amino acids in cell culture (SILAC) (105). The stable cell lines with FL TNK2 or ΔC TNK2 were labeled by growing in heavy (K8, R10) and light (K0, R0) amino acids supplemented SILAC media, respectively, and treated overnight with exogenous biotin **(Figure 4.1D)**. The cells were harvested, lysed, mixed and then digested with trypsin. To identify proteins biotinylated downstream of full-length or truncated TNK2 protein, we employed a method recently developed in our lab that allows the site-specific analysis of biotinylation on peptides via mass spectrometry analysis (106). Using whole cell lysates derived from SILAC-labeled breast cancer cells (HCC1395) expressing either ΔC mutant (Light) or wild-type (Heavy) TNK2-BirA*, we performed immunoaffinity-based enrichment of biotinylated peptides using an anti-biotin antibody.

4.2 Mass spectrometry and data analysis to identify and quantify biotinylated peptides and proteins

As part of our strategy of applying proximity-based biotinylation for mapping the intracellular protein interactors of TNK2, we acquired quantitative mass spectrometry data from the analysis of enriched fractions of biotinylated peptides derived from three

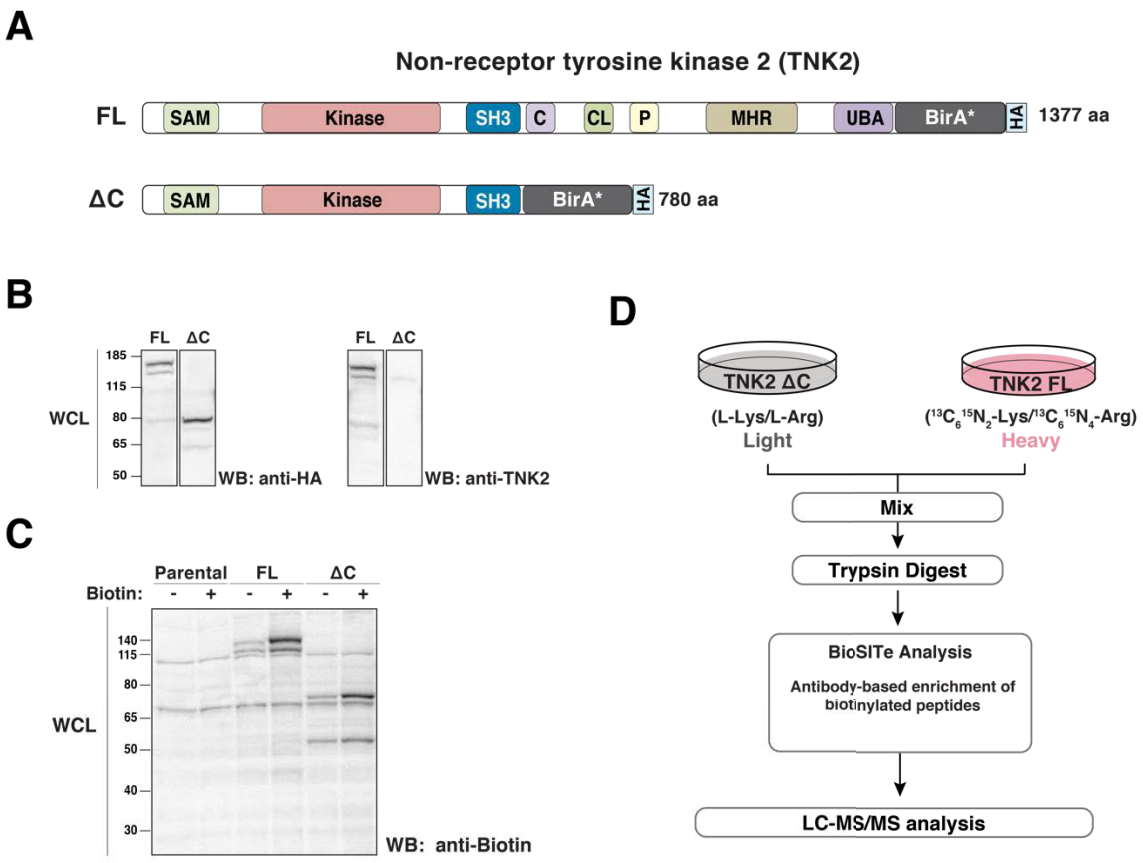

Figure 4.1 Overview of SILAC-BioSITE quantitative workflow for mapping TNK2 interactome in breast cancer cells (A) Schematic depicting the protein domains within recombinant TNK2 full length (top) and ΔC mutant (bottom) constructs used in this study. A biotin ligase (BirA*) and HA tag were cloned in-frame at the C-terminus of both variants as indicated. (B) Western blot analysis of whole cell lysates from breast cancer cells (HCC1395) to confirm expression of TNK2 FL-BirA* or ΔC-BirA*constructs. Analysis was performed using antibodies against TNK2 and HA-tag. (C) Western blot analysis of global biotinylation using whole cell lysates from parental cells and cells expressing TNK2 FL-BirA* or ΔC-BirA*. Analysis was performed using a biotin-specific antibody. (D) Experimental workflow for differential interactome analysis of SILAC-labeled HCC1395 cells expressing full-length (Heavy) or ΔC mutant (Light) TNK2-BirA* fusion constructs were incubated overnight with media containing biotin. Equal amounts of cell lysates from each condition were mixed and digested into peptides. Biotinylated peptides were enriched using BioSITe and analyzed by LC-MS/MS.

biological replicates. LC-MS/MS analysis using an Orbitrap Fusion Lumos generated three raw mass spectrometry files that were then subjected to spectral matching via a combined database search. Our experimental design called for a search scheme that could identify SILAC labeled or unlabeled peptides modified with biotin at one or more/all lysines. To consider all such peptides, we configured our database search parameters to consider the presence of numerous variable modifications, including SILAC modification of lysine and arginine (Lys8, Arg10) biotinylation of light (Lys-Biotin) as well as SILAC-modified lysine (Lys8+Biotin) **(Figure 4.2A)**. We also allowed for 3 missed cleavages to account for the lack of cleavage by trypsin after lysines with biotinylated sidechains. Database spectral matching with the above parameters led to the identification of 688 unique biotinylated peptides corresponding to 367 biotinylated proteins. When possible, MS/MS spectra were manually examined to check for the presence of signature fragment ions, such as those resulting from biotinylated heavy (m/z = 316.17) and light (m/z = 310.15) lysines (107), and to confirm the presence and quality of fragments covering the annotated biotinylation site.

The list of biotinylated peptides identified from the coupled use of SILAC-labeling with BioSITe called for an adequate quantitation platform. As depicted in **Figure 4.2A**, the biotinylated peptides in our study could be present in multiple forms and could be isotopically-labeled (Lys8), biotinylated, or both. With the added consideration of miscleaved peptides, these parameters collectively lead to a massively large search space, resulting in failure to quantify certain peptides with multiple dynamic modifications, making quantification of peptides with multiply modified lysines quite challenging. Using a quantification workflow that integrated PyQuant, a quantification algorithm previously

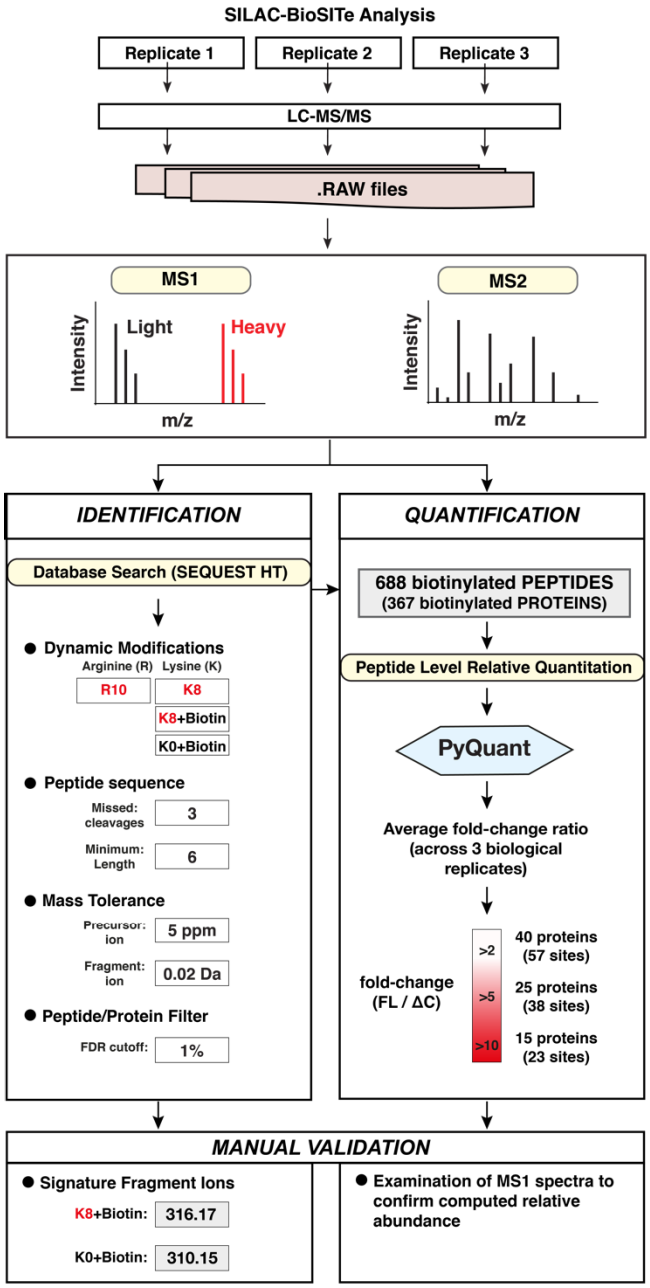

Figure 4.2 Computational approach to identify and quantify biotinylated peptides. The in-silico pipeline used for processing raw files to identify and quantify biotinylated peptides is depicted. Raw files were processed to generate a list of biotinylated peptide IDs. (A) Workflow used to identify biotinylated peptides sequences and proteins. Relevant details regarding parameters used for the automated database search, such as dynamic modification and mass tolerance thresholds are indicated. Strategy used when manually examining spectra is also depicted. The workflow used to quantify the identified biotinylated peptides is also summarized. Quantification was performed by configuring the pipeline offered by PyQuant as described previously (1).

developed by our group (108), we derived MS1 level quantification for the list of identified biotinylated species and their corresponding isotopic counterparts **(Figure 4.2A)**. This enabled us to get quantitation for a large majority of the identified biotinylated peptides and proteins. As with the results from database identification, we manually examined MS1 spectra for a selection of biotinylated candidates to confirm the validity of the relative abundance as reported by PyQuant. Representative MS2 along with the corresponding MS1 spectrum used for deriving relative quantification are shown for biotinylated peptides from NCK interactor 1 (NCK1) **(Figure 4.3A)**, Cortactin (CTTN) **(Figure 4.3B)**, Clathrin interactor 1 (CLINT1) (Figure 4.3C) and TELO2-interactor 1 (TTI1) protein **(Figure 4.3D)**.

4.3 Identification of TNK2 proximal interactome using BioSITe

Our group has previously shown that in the analysis of data derived from a BioSITe study, the degree of protein biotinylation can be used as a potential metric to help prioritize confident protein interactors from potential contaminants and false positives (106). Therefore, we grouped the biotinylated proteins in our data based on their degree of biotinylation **(Figure 4.4A)**. As expected, this resulted in the sorting of TNK2 bait protein as the protein showing the highest degree of biotinylation. TNK2 is known to bind to itself and form homo-dimers (109). Therefore, in addition to autologous biotinylation, identification of biotinylated peptides mapping to TNK2 may in part be a result of TNK2 self-dimerization. However, we did not observe many of the well-established direct binders of TNK2 among the proteins showing the highest degree of biotinylation. In search for true interaction partners of TNK2, we next examined the set of biotinylated sites and proteins

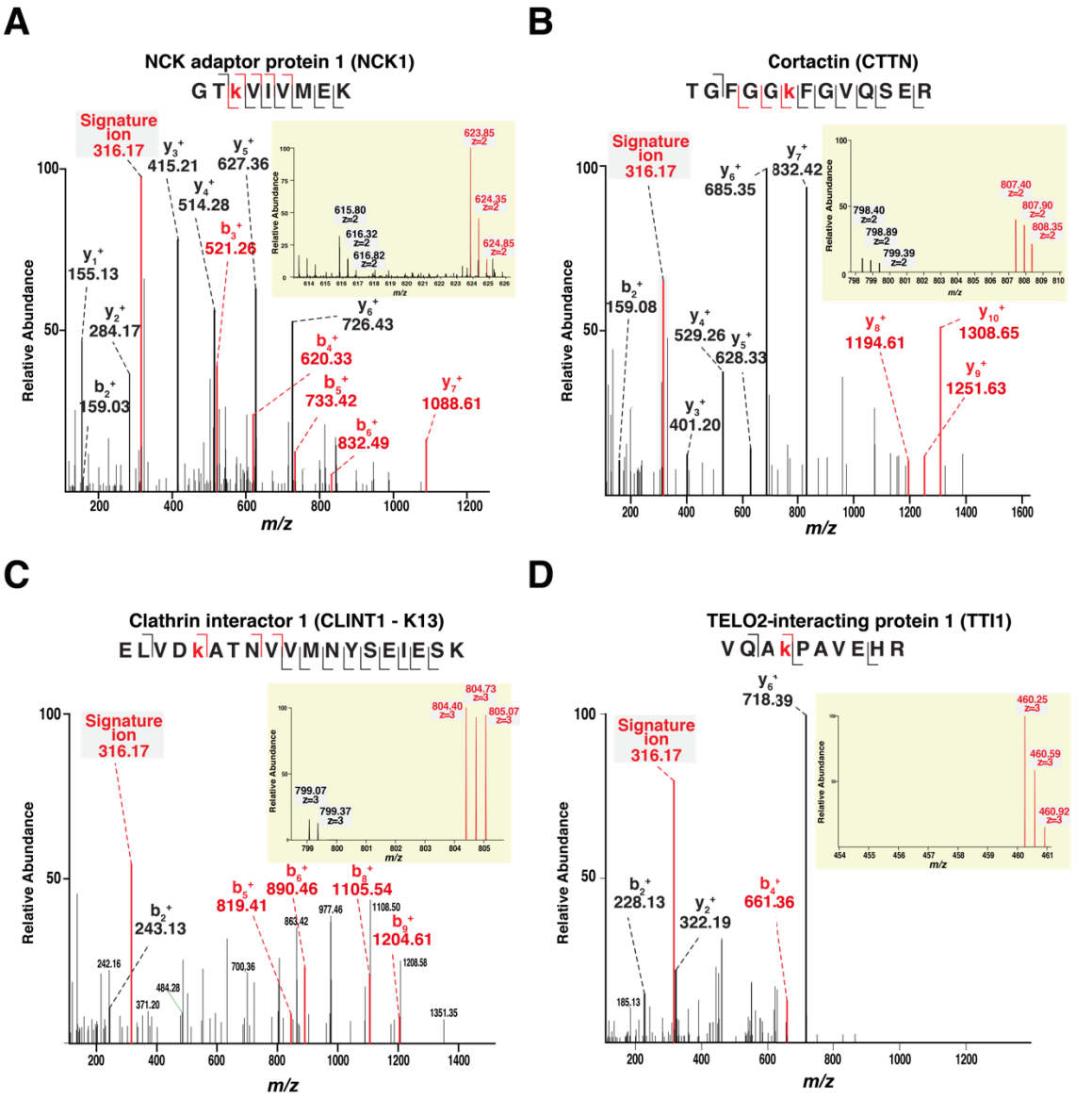

Figure 4.3 Representative MS/MS and MS spectra used for identification and relative quantitation of biotinylated peptides derived from cells expressing TNK2-BirA* constructs. (A) MS/MS fragmentation of a biotinylated peptide mapping to NCK adaptor protein 1 (NCK1), a known binding partner of TNK2. Representative MS1 spectrum used for deriving relative quantification between full-length and ΔC mutant cells is shown as an inset. Fragments ion that confirm the annotated site of biotinylation are shown in red. Signature fragment ions resulting from biotinylated lysine moieties are indicated in green. (B) Representative MS/MS and MS1 spectra of biotinylated peptides mapping to K272 in cortactin (CTTN), a known interactor and phosphorylation substrate of TNK2. (C) Representative MS2 and MS spectra of biotinylated peptides mapping to selected novel TNK2 interactors such as (C) clathrin interactor 1(CLINT1) and (D) TELO2-interacting protein 1 (TTI1).

that showed a significantly higher abundance in cells expressing full-length TNK2 relative to cells expressing a truncated version of the protein. This led to the identification of 57 biotinylation sites mapping to 40 biotinylated proteins that displayed a ≥2-fold higher abundance in cells expressing full-length relative to truncated protein (referred to from here on as full length specific TNK2 binders). Among these were biotinylation sites mapping to several experimentally established TNK2 binders, including NCK adaptor protein 1 (NCK1), NCK adaptor protein 2 (NCK2) (110,111), cortactin (CTTN) (112), clathrin (CLTC) (113), and Cyclin G associated kinase (GAK). Some of our biotinylated candidates overlapped with those identified in previous co-immunoprecipitation experiments using recombinant TNK2, including STAT3 (K707) and CRKL (K89). Analysis in the previous study also showed that TNK2 was involved in phosphorylation of STAT3 at Y705, and that this process is partly regulated by HSP90 (114). In line with these reports, we also observed increased biotinylation of several heat shock protein complex components in the cells expressing full-length TNK2-BirA*, including HSP8 (K507, K512), HSP90AB1 (K607), and HSPA1A (K507). Furthermore, while CRKL was identified in the above study, we also found biotinylated peptides mapping to CRK, which is not a known TNK2 binder. Overall, these findings indicate that the relative abundance ratio for biotinylation sites/peptides between full-length and truncated constructs can serve as a powerful differentiator between true binders of the biotin ligase fused bait protein and background or false positives. In this study, this strategy allowed us to identify true binders of TNK2 protein by relative comparison of protein biotinylation between cells expressing full-length TNK2 and truncated TNK2.

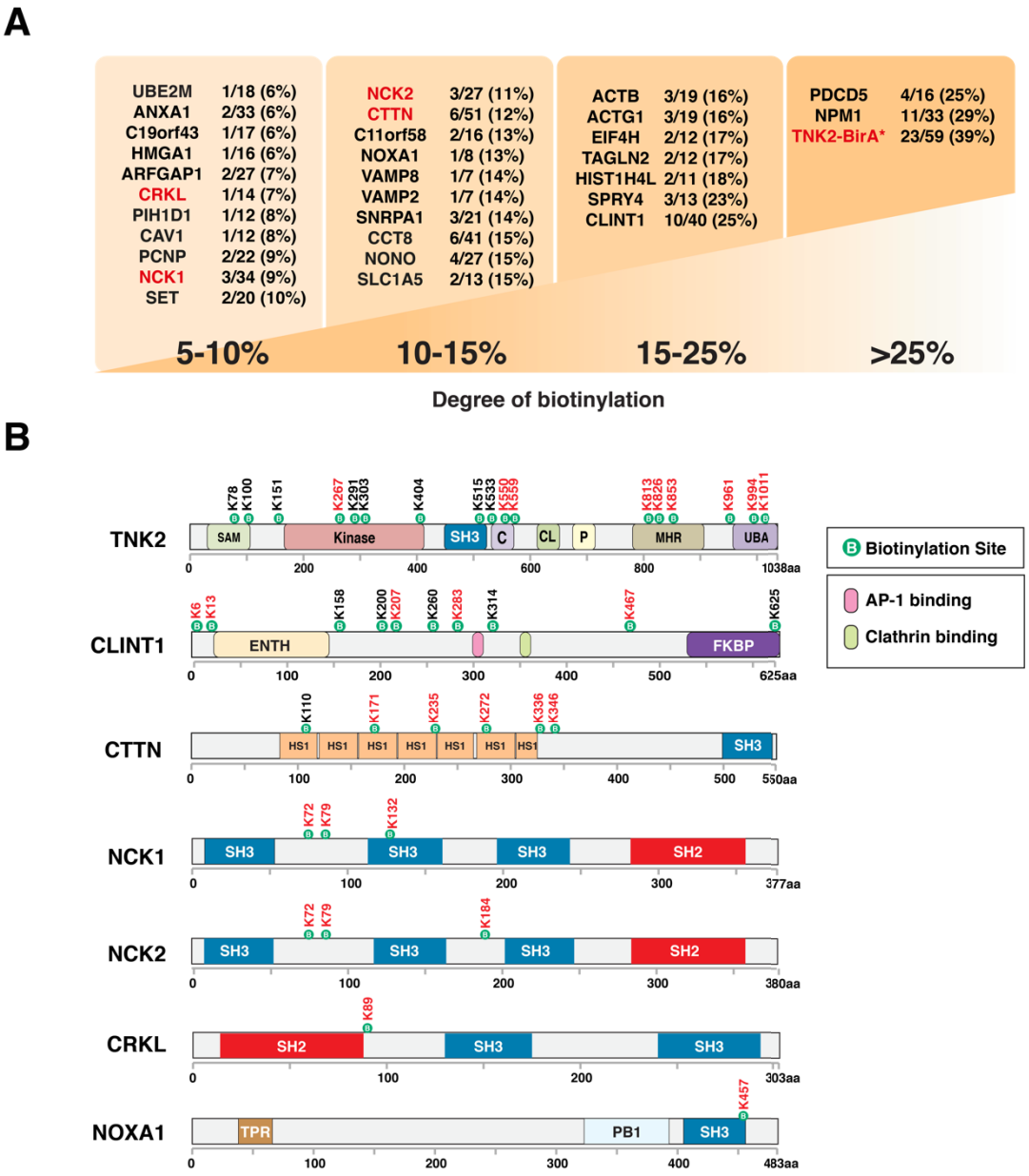

Figure 4.4. Mapping of TNK2 interactome using biotinylation sites. (A) Biotinylated proteins identified by BioSITe are shown grouped by the degree of biotinylation in cells expressing TNK2 constructs fused with biotin ligase. The number of biotinylated proteins falling within each group are indicated. For each group, 10 representative biotinylated proteins are shown. The number of biotinylated lysines out of the total number of lysines are provided for each protein. Known binding partners of TNK2 are shown in red. (B) Domain schematics for select biotinylated proteins identified in this study. The protein domains and the relative localization of the biotinylation sites mapped to the protein are shown. The full names of the protein domains are described in the legend. Sites showing a 2-fold or greater change in abundance ratio between TNK2 FL-BirA* and TNK2 ΔC-BirA* are labeled in red.

In addition to identifying sites of biotinylation on several proteins that overlap with published interactors and substrates of TNK2, our BioSITe data also included many biotinylated proteins that have not been previously reported to be associated with TNK2 and could be potentially novel TNK2 interactors. These interaction candidates included clathrin interactor 1 (CLINT1), TELO2 interacting protein 1 (TTI1), NADPH oxidase activator 1 (NOXA1), RuvB like AAA ATPase 1 (RUVBL1), and GRB10 interacting GYF protein 2 (GIGYF2). We noted that several of the novel interaction candidates for TNK2 are proteins that bind known interactors of TNK2, as determined by previously reported studies. For instance, comparison of 101 known CLINT1 interactors with known TNK2 interactors indicates common binding to several well-characterized TNK2 interactors, included CLTC, GAK, NCK1, GTSE1, and NTRK1. Similarly, in the database derived interactome for TTI1 (n = 47), an important member of the mTORC protein complexes (115), binds EGFR and NTRK1.

TNK2 is known to bind SH3 domains through its proline rich domain (116–118). To test for an enriched presence of SH3 and potentially other protein domains that mediated molecular association with proteins, we performed analyses to determine whether proteins showing full-length specific biotinylation were enriched in certain protein domains, and also whether the site of biotinylation in these putative TNK2 binding partners were preferentially localized to specific protein domains and regions. A more in-depth examination of biotinylation site localization in the context of the annotated domains in individual proteins revealed that many sites mapped to known protein domains and motifs, namely SH3 domains (CRK, CTTN, DNMBP, NCK1, NCK2, NOXA1, STAM2, TNK2, UBASH3B) and the ENTH domain which is involved primarily in mediating binding to

membrane lipids (CLINT1, HGS, HIP1, STAM2, TTI1). A few selected biotinylated proteins and the location of biotinylation sites with their linear protein structure is depicted in **Figure 4.4B**. Furthermore, we observed that while several proteins in our data were found biotinylated at multiple sites, the relative abundance of the biotinylated peptides mapping to regions adjacent to these protein domains were often more abundant in cells expressing full-length TNK2 where the proline-rich domain is conserved.

4.4 TNK2 proximally-biotinylated proteins are involved in vesicle-mediated endocytosis

A collective survey of previous studies available on TNK2 indicates its association with pathways involved in recycling of surface receptors, especially receptor tyrosine kinases such as Epidermal growth factor receptor (EGFR), Insulin receptor, and AXL receptor tyrosine kinase (119–121). However, the specific proteins and protein complexes engaged by TNK2 to drive these processes have not been properly surveyed previously. Thus, we performed GO enrichment analysis to determine whether proteins showing full-length specific biotinylation where known to be involved with these and other novel biological pathways and processes **(Figure 4.5A)**. Survey of biological processes and KEGG pathways revealed an enrichment of proteins mapping to pathways associated with either endocytosis (CLINT1, CLTC, CTTN, GAK, HIP1, HSP90AA1, HSPA8), receptor tyrosine kinase signaling (GIGYF2, HGS, HSP90AA1, STAM2, STAT3), or both (CRK, EPS15L1, NCK1, TNK2) **(Figure 4.5A)**. Next, we sought to determine whether there was a prevalence of known protein-protein interactions between the full-length TNK2-specific biotinylation proteins identified in our dataset. To identify such interactions, we performed

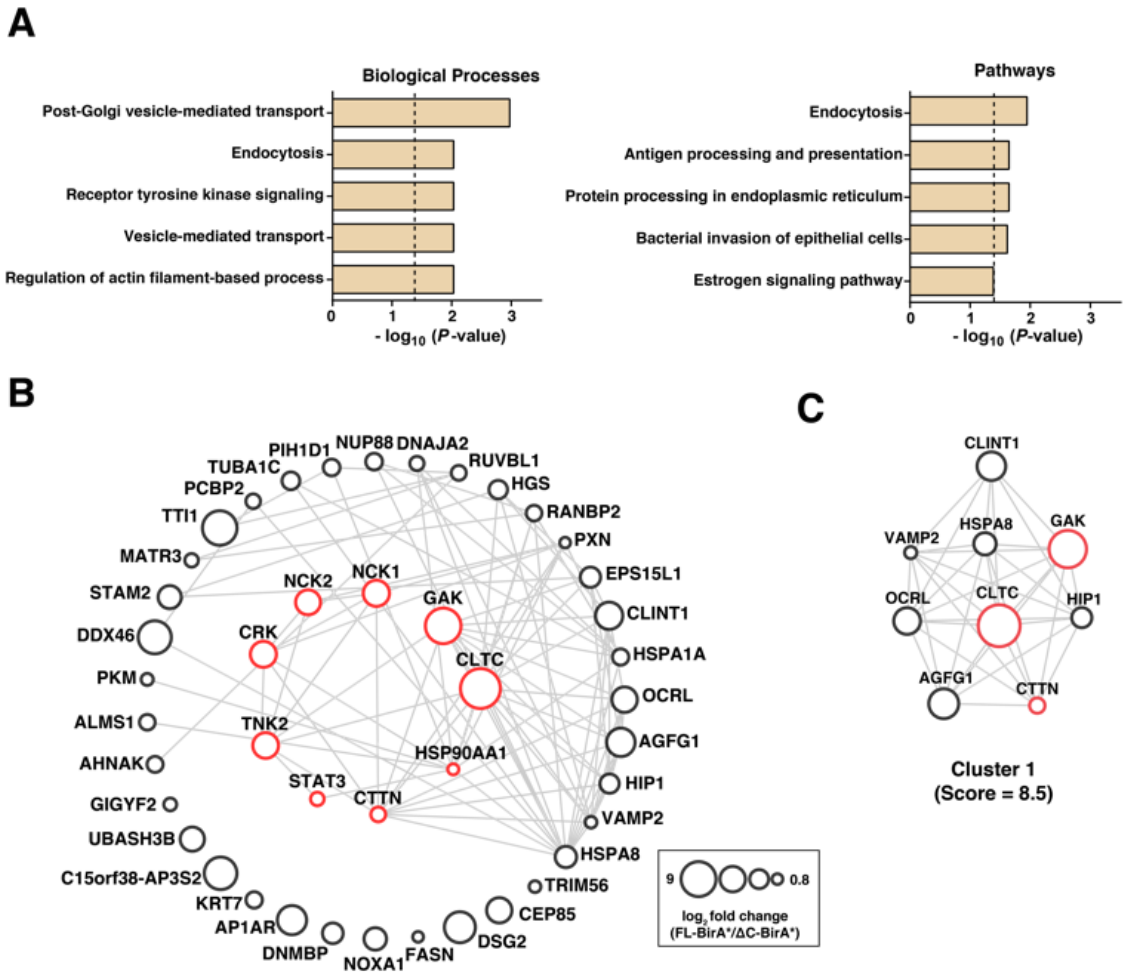

Figure 4.5. Pathway enrichment and STRING analysis of protein significantly biotinylated in cells expressing TNK2 FL-BirA* construct. (A) Results from Gene Ontology (GO) enrichment analysis of biotinylated proteins that displayed a ≥2-fold higher abundance in cells expressing full-length relative to truncated protein (n = 40). Top five biological processes and KEGG pathways found to be significantly enriched ($p < 0.05$) are shown. Dashed line denotes p-value threshold for 0.05. (B) Results from STRING analysis of biotinylated proteins that displayed a ≥2-fold higher abundance in cells expressing full-length relative to truncated protein. The image was created using Cytoscape. The node size is proportional to the protein level fold-change (FL-BirA* vs. ΔC-BirA*) for each biotinylated protein, with the smallest circle corresponding to a 2-fold change and the largest circle corresponding to a 20-fold change. Known interactors or phosphorylation substrates of TNK2 based on database and literature searches are labeled in red. (C) Top scoring cluster from the protein interaction network shown in (B). Identification and analysis of enriched sub-clusters was conducted using the MCODE clustering algorithm.

STRING analysis, which revealed the presence of experimental verified protein-protein interactions amongst the proteins found to show significant biotinylation in full-length TNK2 **(Figure 4.5B)**. Within the resulting protein interaction network, we were able to identify several functional clusters or interaction modules, potentially indicating proximal protein complexes. Members of the top scoring cluster **(Figure 4.5C)** included already known TNK2 interactors, including CTTN, CLTC, and GAK, as well as several novel interaction candidates, such as AGFG1, CLINT1, HIP1, HSP8, OCRL, and VAMP2. This sub-cluster accounted for a majority of the proteins that were found in our GO enrichment analysis to be associated with endocytosis and related processes, indicating a potential protein complex that may be involved in carrying out TNK2-mediated endocytosis. We were able to validate interaction between TNK2 and a member of this sub-cluster, Clathrin interactor 1 (CLINT1), by performing immunoblot analysis on proteins that co-immunoprecipitated (co-IP) from an anti-HA tag pulldown **(Figure 4.6B).** CLINT1 was observed only in eluates obtained from anti-HA tag IP from lysates of cells expressing TNK2 FL-BirA*, suggesting this interaction required the full-length TNK2 protein. As a positive control, we performed a parallel co-IP for confirming interaction between full-length TNK2 and NCK1, a well-established binding partner for TNK2 **(Figure 4.6A).** Therefore, our analysis demonstrates that a differential interactome study using a SILAC-BioSITe approach can enable the detection and identification of true novel interactors of a protein of interest.

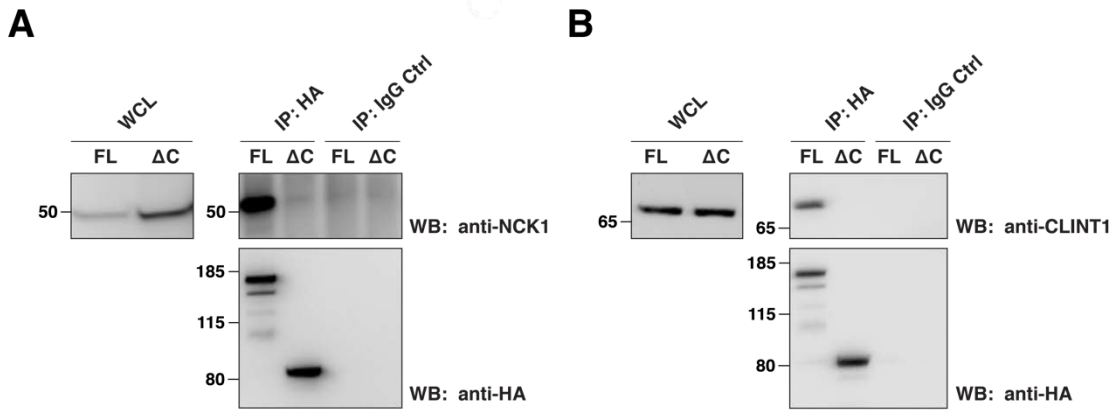

Figure 4.6. Validation of CLINT1 interaction with full-length TNK2-BirA* using co-immunoprecipitation analysis. (A) Western blot analysis of control mouse IgG and anti-HA tag IP from cells expressing TNK2 FL-BirA* or ΔC-BirA*. Western blot analysis was performed using antibodies against NCK adaptor protein 1 (NCK1) and HA-tag. Analysis of whole cell lysates with anti-NCK1 was performed in parallel. (B) Western blot analysis of control mouse IgG and anti-HA tag IP from cells expressing TNK2 FL-BirA* or ΔC-BirA*. Western blot analysis was performed using antibodies against Clathrin interactor 1 (CLINT1) and HA-tag. Analysis of whole cell lysates with anti-CLINT1 was performed in parallel.

4.5 Conclusions

Large-scale interaction studies employing high-throughput platforms have become increasing popular for obtaining "comprehensive" protein-protein interactome maps of protein families (122) and even entire human interactomes (123,124). However, there is a need for appropriate controls or complementary methods that help pinpoint true interactors among the large list of candidates rendered from one method alone. BioID and other proximity based labeling strategies can serve as good complementary methods to assist in the prioritization and sorting of high confidence protein interactors. In fact, large scale studies based on proximity dependent biotinylation are already becoming popular for mapping interactomes for protein families, including G protein coupled receptors (125,126), protein phosphatases (127), nuclear transport receptors (128), as well as for individual proteins like paxillin (129), fibroblast growth factor receptors (FGFRs) (130,131), transcription factor SOX2 (132), dynein (133), and many others. However, these methods are not being applied for studying molecules whose protein interactomes have not yet been comprehensively characterized. In the present work, we described how we mapped the interactome of one of the few understudied but important kinases: non-receptor tyrosine kinase 2 (TNK2). TNK2 has an established role in driving various cancers, and we also recently identified it as one of the proteins that exhibits increased tyrosine hyperphosphorylation downstream of two different oncogenic KRAS mutations. We show the coupled use of BioID with BioSITe, which allowed us to obtain site-level biotinylation evidence for putative TNK2 binders. Furthermore, we show how to use this strategy for a quantitative proteomics experiment employing SILAC to compare the relative abundance of biotinylated proteins between full-length and truncated version of TNK2. This approach allowed us to recapitulate known interactors of TNK2, but also identify several novel TNK2

interactors, including clathrin interactor 1 (CLINT1). Further experiments are warranted to assess the functional significance of novel TNK2 interactors in TNK2-mediated signaling. We believe that the strategy used in this work is broadly applicable for the field of protein interactomes and should be applied to map the interactomes of other proteins of interest.

BIBLIOGRAPHY

1. WHO | World Health Statistics 2018: Monitoring health for the SDGs [Internet]. Available from:http://www.who.int/gho/publications/world_health_statistics/2018/en/

2. Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med. 2017 Jun;23(6):703–13.

3. Simanshu DK, Nissley DV, McCormick F. RAS Proteins and Their Regulators in Human Disease. Cell. 2017 Jun 29;170(1):17–33.

4. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012 May;2(5):401–4.

5. Almoguera C, Shibata D, Forrester K, Martin J, Arnheim N, Perucho M. Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. Cell. 1988 May 20;53(4):549–54.

6. Prior IA, Lewis PD, Mattos C. A comprehensive survey of Ras mutations in cancer. Cancer Res. 2012 May 15;72(10):2457–67.

7. Cherfils J, Chardin P. GEFs: structural basis for their activation of small GTP-binding proteins. Trends Biochem Sci. 1999 Aug;24(8):306–11.

8. Hunter JC, Manandhar A, Carrasco MA, Gurbani D, Gondi S, Westover KD. Biochemical and Structural Analysis of Common Cancer-Associated KRAS Mutations. Mol Cancer Res MCR. 2015 Sep;13(9):1325–35.

9. Smith MJ, Neel BG, Ikura M. NMR-based functional profiling of RASopathies and oncogenic RAS mutations. Proc Natl Acad Sci U S A. 2013 Mar 19;110(12):4574–9.

10. Shao DD, Xue W, Krall EB, Bhutkar A, Piccioni F, Wang X, et al. KRAS and YAP1 converge to regulate EMT and tumor survival. Cell. 2014 Jul 3;158(1):171–84.

11. Zhang W, Nandakumar N, Shi Y, Manzano M, Smith A, Graham G, et al. Downstream of mutant KRAS, the transcription regulator YAP is essential for neoplastic progression to pancreatic ductal adenocarcinoma. Sci Signal. 2014 May 6;7(324):ra42.

12. Nussinov R, Tsai C-J, Jang H, Korcsmáros T, Csermely P. Oncogenic KRAS signaling and YAP1/β-catenin: Similar cell cycle control in tumor initiation. Semin Cell Dev Biol. 2016;58:79–85.

13. Nussinov R, Muratcioglu S, Tsai C-J, Jang H, Gursoy A, Keskin O. The Key Role of Calmodulin in KRAS-Driven Adenocarcinomas. Mol Cancer Res MCR. 2015 Sep;13(9):1265–73.

14. Nussinov R, Wang G, Tsai C-J, Jang H, Lu S, Banerjee A, et al. Calmodulin and PI3K Signaling in KRAS Cancers. Trends Cancer. 2017;3(3):214–24.

15. Kawada K, Toda K, Sakai Y. Targeting metabolic reprogramming in KRAS-driven cancers. Int J Clin Oncol. 2017 Aug;22(4):651–9.

16. Toda K, Kawada K, Iwamoto M, Inamoto S, Sasazuki T, Shirasawa S, et al. Metabolic Alterations Caused by KRAS Mutations in Colorectal Cancer Contribute to Cell Adaptation to Glutamine Depletion by Upregulation of Asparagine Synthetase. Neoplasia N Y N. 2016 Nov;18(11):654–65.

17. Kerr EM, Martins CP. Metabolic rewiring in mutant Kras lung cancer. FEBS J. 2018 Jan;285(1):28–41.

18. Weinberg F, Hamanaka R, Wheaton WW, Weinberg S, Joseph J, Lopez M, et al. Mitochondrial metabolism and ROS generation are essential for Kras-mediated tumorigenicity. Proc Natl Acad Sci. 2010 May 11;107(19):8788–93.

19. Kerr EM, Gaude E, Turrell FK, Frezza C, Martins CP. Mutant Kras copy number defines metabolic reprogramming and therapeutic susceptibilities. Nature. 2016 Mar 3;531(7592):110–3.

20. Ying H, Kimmelman AC, Lyssiotis CA, Hua S, Chu GC, Fletcher-Sananikone E, et al. Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. Cell. 2012 Apr 27;149(3):656–70.

21. Son J, Lyssiotis CA, Ying H, Wang X, Hua S, Ligorio M, et al. Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. Nature. 2013 Apr 4;496(7443):101–5.

22. Cheong H, Lu C, Lindsten T, Thompson CB. Therapeutic targets in cancer cell metabolism and autophagy. Nat Biotechnol. 2012 Jul 10;30(7):671–8.

23. Kim M-J, Woo S-J, Yoon C-H, Lee J-S, An S, Choi Y-H, et al. Involvement of autophagy in oncogenic K-Ras-induced malignant cell transformation. J Biol Chem. 2011 Apr 15;286(15):12924–32.

24. Padanad MS, Konstantinidou G, Venkateswaran N, Melegari M, Rindhe S, Mitsche M, et al. Fatty Acid Oxidation Mediated by Acyl-CoA Synthetase Long Chain 3 Is Required for Mutant KRAS Lung Tumorigenesis. Cell Rep. 2016 09;16(6):1614–28.

25. Cox AD, Fesik SW, Kimmelman AC, Luo J, Der CJ. Drugging the undruggable RAS: Mission possible? Nat Rev Drug Discov. 2014 Nov;13(11):828–51.

26. Wang W, Fang G, Rudolph J. Ras inhibition via direct Ras binding--is there a path forward? Bioorg Med Chem Lett. 2012 Sep 15;22(18):5766–76.

27. Zeitouni D, Pylayeva-Gupta Y, Der CJ, Bryant KL. KRAS Mutant Pancreatic Cancer:

No Lone Path to an Effective Treatment. Cancers. 2016 Apr 18;8(4).

28. Gibbs JB, Graham SL, Hartman GD, Koblan KS, Kohl NE, Omer CA, et al. Farnesyltransferase inhibitors versus Ras inhibitors. Curr Opin Chem Biol. 1997 Aug;1(2):197–203.

29. Bergo MO, Gavino BJ, Hong C, Beigneux AP, McMahon M, Casey PJ, et al. Inactivation of Icmt inhibits transformation by oncogenic K-Ras and B-Raf. J Clin Invest. 2004 Feb;113(4):539–50.

30. Spiegel J, Cromm PM, Zimmermann G, Grossmann TN, Waldmann H. Small-molecule modulation of Ras signaling. Nat Chem Biol. 2014 Aug;10(8):613–22.

31. Young A, Lyons J, Miller AL, Phan VT, Alarcón IR, McCormick F. Ras signaling and therapies. Adv Cancer Res. 2009;102:1–17.

32. Gysin S, Salt M, Young A, McCormick F. Therapeutic strategies for targeting ras proteins. Genes Cancer. 2011 Mar;2(3):359–72.

33. Chuang H-C, Huang P-H, Kulp SK, Chen C-S. Pharmacological strategies to target oncogenic KRAS signaling in pancreatic cancer. Pharmacol Res. 2017 Mar 1;117:370–6.

34. Lito P, Solomon M, Li L-S, Hansen R, Rosen N. Allele-specific inhibitors inactivate mutant KRAS G12C by a trapping mechanism. Science. 2016 Feb 5;351(6273):604–8.

35. Zeng M, Lu J, Li L, Feru F, Quan C, Gero TW, et al. Potent and Selective Covalent Quinazoline Inhibitors of KRAS G12C. Cell Chem Biol. 2017 Aug 17;24(8):1005-1016.e3.

36. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009 Nov 5;462(7269):108–12.

37. Marcotte R, Brown KR, Suarez F, Sayad A, Karamboulas K, Krzyzanowski PM, et al. Essential gene profiles in breast, pancreatic, and ovarian cancer cells. Cancer Discov. 2012 Feb;2(2):172–89.

38. Steckel M, Molina-Arcas M, Weigelt B, Marani M, Warne PH, Kuznetsov H, et al. Determination of synthetic lethal interactions in KRAS oncogene-dependent cancer cells reveals novel therapeutic targeting strategies. Cell Res. 2012 Aug;22(8):1227–45.

39. Luo J, Emanuele MJ, Li D, Creighton CJ, Schlabach MR, Westbrook TF, et al. A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. Cell. 2009 May 29;137(5):835–48.

40. Bridgett S, Campbell J, Lord CJ, Ryan CJ. CancerGD: A Resource for Identifying and

Interpreting Genetic Dependencies in Cancer. Cell Syst. 2017 Jul 26;5(1):82-86.e3.

41. Martin TD, Cook DR, Choi MY, Li MZ, Haigis KM, Elledge SJ. A Role for Mitochondrial Translation in Promotion of Viability in K-Ras Mutant Cells. Cell Rep. 2017 Jul 11;20(2):427–38.

42. Yau EH, Kummetha IR, Lichinchi G, Tang R, Zhang Y, Rana TM. Genome-Wide CRISPR Screen for Essential Cell Growth Mediators in Mutant KRAS Colorectal Cancers. Cancer Res. 2017 15;77(22):6330–9.

43. Wang J, Hu K, Guo J, Cheng F, Lv J, Jiang W, et al. Suppression of KRas-mutant cancer through the combined inhibition of KRAS with PLK1 and ROCK. Nat Commun. 2016 May 19;7:11363.

44. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009 Nov 5;462(7269):108–12.

45. Weng M-T, Lee J-H, Wei S-C, Li Q, Shahamatdar S, Hsu D, et al. Evolutionarily conserved protein ERH controls CENP-E mRNA splicing and is required for the survival of KRAS mutant cancer cells. Proc Natl Acad Sci U S A. 2012 Dec 26;109(52):E3659-3667.

46. Scholl C, Fröhling S, Dunn IF, Schinzel AC, Barbie DA, Kim SY, et al. Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. Cell. 2009 May 29;137(5):821–34.

47. Narvaez AJ, Ber S, Crooks A, Emery A, Hardwick B, Guarino Almeida E, et al. Modulating Protein-Protein Interactions of the Mitotic Polo-like Kinases to Target Mutant KRAS. Cell Chem Biol. 2017 Aug 17;24(8):1017-1028.e7.

48. Benvenuti S, Sartore-Bianchi A, Di Nicolantonio F, Zanon C, Moroni M, Veronese S, et al. Oncogenic activation of the RAS/RAF signaling pathway impairs the response of metastatic colorectal cancers to anti-epidermal growth factor receptor antibody therapies. Cancer Res. 2007 Mar 15;67(6):2643–8.

49. Ramos FJ, Macarulla T, Capdevila J, Elez E, Tabernero J. Understanding the predictive role of K-ras for epidermal growth factor receptor-targeted therapies in colorectal cancer. Clin Colorectal Cancer. 2008 Dec;7 Suppl 2:S52-57.

50. Allegra CJ, Jessup JM, Somerfield MR, Hamilton SR, Hammond EH, Hayes DF, et al. American Society of Clinical Oncology provisional clinical opinion: testing for KRAS gene mutations in patients with metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy. J Clin Oncol Off J Am Soc Clin Oncol. 2009 Apr 20;27(12):2091–6.

51. Lièvre A, Bachet J-B, Le Corre D, Boige V, Landi B, Emile J-F, et al. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. Cancer Res.

2006 Apr 15;66(8):3992–5.

52. Normanno N, Tejpar S, Morgillo F, De Luca A, Van Cutsem E, Ciardiello F. Implications for KRAS status and EGFR-targeted therapies in metastatic CRC. Nat Rev Clin Oncol. 2009 Sep;6(9):519–27.

53. Misale S, Yaeger R, Hobor S, Scala E, Janakiraman M, Liska D, et al. Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. Nature. 2012 Jun 28;486(7404):532–6.

54. Cepero V, Sierra JR, Corso S, Ghiso E, Casorzo L, Perera T, et al. MET and KRAS gene amplification mediates acquired resistance to MET tyrosine kinase inhibitors. Cancer Res. 2010 Oct 1;70(19):7580–90.

55. Matsusaka S, Kobunai T, Yamamoto N, Chin K, Ogura M, Tanaka G, et al. Prognostic impact of KRAS mutant type and MET amplification in metastatic and recurrent gastric cancer patients treated with first-line S-1 plus cisplatin chemotherapy. Genes Cancer. 2016 Jan;7(1–2):27–35.

56. Sartore-Bianchi A, Valtorta E, Amatu A, Veronese S, Lauricella C, Bonazzina E, et al. Clonal evolution and KRAS-MET coamplification during secondary resistance to EGFR-targeted therapy in metastatic colorectal cancer. ESMO Open. 2016;1(4):e000079.

57. Renaud S, Guerrera F, Seitlinger J, Costardi L, Schaeffer M, Romain B, et al. KRAS exon 2 codon 13 mutation is associated with a better prognosis than codon 12 mutation following lung metastasectomy in colorectal cancer. Oncotarget. 2016 Nov 29;

58. Messner I, Cadeddu G, Huckenbeck W, Knowles HJ, Gabbert HE, Baldus SE, et al. KRAS p.G13D mutations are associated with sensitivity to anti-EGFR antibody treatment in colorectal cancer cell lines. J Cancer Res Clin Oncol. 2013 Feb;139(2):201–9.

59. Kumar SS, Price TJ, Mohyieldin O, Borg M, Townsend A, Hardingham JE. KRAS G13D Mutation and Sensitivity to Cetuximab or Panitumumab in a Colorectal Cancer Cell Line Model. Gastrointest Cancer Res GCR. 2014 Jan;7(1):23–6.

60. Park JT, Johnson N, Liu S, Levesque M, Wang YJ, Ho H, et al. Differential in vivo tumorigenicity of diverse KRAS mutations in vertebrate pancreas: A comprehensive survey. Oncogene. 2015 May 21;34(21):2801–6.

61. Jones RP, Sutton PA, Evans JP, Clifford R, McAvoy A, Lewis J, et al. Specific mutations in KRAS codon 12 are associated with worse overall survival in patients with advanced and recurrent colorectal cancer. Br J Cancer. 2017 Mar 28;116(7):923–9.

62. Rowland A, Dias MM, Wiese MD, Kichenadasse G, McKinnon RA, Karapetis CS, et al. Meta-analysis comparing the efficacy of anti-EGFR monoclonal antibody therapy

between KRAS G13D and other KRAS mutant metastatic colorectal cancer tumours. Eur J Cancer Oxf Engl 1990. 2016 Mar;55:122–30.

63. Guerrero S, Casanova I, Farré L, Mazo A, Capellà G, Mangues R. K-ras codon 12 mutation induces higher level of resistance to apoptosis and predisposition to anchorage-independent growth than codon 13 mutation or proto-oncogene overexpression. Cancer Res. 2000 Dec 1;60(23):6750–6.

64. Alamo P, Gallardo A, Di Nicolantonio F, Pavón MA, Casanova I, Trias M, et al. Higher metastatic efficiency of KRas G12V than KRas G13D in a colorectal cancer model. FASEB J Off Publ Fed Am Soc Exp Biol. 2015 Feb;29(2):464–76.

65. Conzelmann M, Linnemann U, Berger MR. K-ras codon 12 and 13 mutations are correlated with differential patterns of tumor cell dissemination in colorectal cancer patients. Int J Oncol. 2004 Jun;24(6):1537–44.

66. Vizan P, Boros LG, Figueras A, Capella G, Mangues R, Bassilian S, et al. K-ras codon-specific mutations produce distinctive metabolic phenotypes in NIH3T3 mice [corrected] fibroblasts. Cancer Res. 2005 Jul 1;65(13):5512–5.

67. Blons H, Emile JF, Le Malicot K, Julié C, Zaanan A, Tabernero J, et al. Prognostic value of KRAS mutations in stage III colon cancer: post hoc analysis of the PETACC8 phase III trial dataset. Ann Oncol Off J Eur Soc Med Oncol. 2014 Dec;25(12):2378–85.

68. Imamura Y, Morikawa T, Liao X, Lochhead P, Kuchiba A, Yamauchi M, et al. Specific mutations in KRAS codons 12 and 13, and patient prognosis in 1075 BRAF wild-type colorectal cancers. Clin Cancer Res Off J Am Assoc Cancer Res. 2012 Sep 1;18(17):4753–63.

69. Ye X, Chan KC, Waters AM, Bess M, Harned A, Wei B-R, et al. Comparative proteomics of a model MCF10A-KRasG12V cell line reveals a distinct molecular signature of the KRasG12V cell surface. Oncotarget. 2016 Nov 24;

70. Stolze B, Reinhart S, Bulllinger L, Fröhling S, Scholl C. Comparative analysis of KRAS codon 12, 13, 18, 61, and 117 mutations using human MCF10A isogenic cell lines. Sci Rep. 2015 Feb 23;5:8535.

71. Haley JA, Haughney E, Ullman E, Bean J, Haley JD, Fink MY. Altered Transcriptional Control Networks with Trans-Differentiation of Isogenic Mutant-KRas NSCLC Models. Front Oncol. 2014;4:344.

72. Modest DP, Camaj P, Heinemann V, Schwarz B, Jung A, Laubender RP, et al. KRAS allel-specific activity of sunitinib in an isogenic disease model of colorectal cancer. J Cancer Res Clin Oncol. 2013 Jun;139(6):953–61.

73. Demory Beckler M, Higginbotham JN, Franklin JL, Ham A-J, Halvey PJ, Imasuen IE, et al. Proteomic analysis of exosomes from mutant KRAS colon cancer cells identifies

intercellular transfer of mutant KRAS. Mol Cell Proteomics MCP. 2013 Feb;12(2):343–55.

74. Guha U, Chaerkady R, Marimuthu A, Patterson AS, Kashyap MK, Harsha HC, et al. Comparisons of tyrosine phosphorylated proteins in cells expressing lung cancer-specific alleles of EGFR and KRAS. Proc Natl Acad Sci U S A. 2008 Sep 16;105(37):14112–7.

75. Vartanian S, Bentley C, Brauer MJ, Li L, Shirasawa S, Sasazuki T, et al. Identification of mutant K-Ras-dependent phenotypes using a panel of isogenic cell lines. J Biol Chem. 2013 Jan 25;288(4):2403–13.

76. Brunelli L, Caiola E, Marabese M, Broggini M, Pastorelli R. Comparative metabolomics profiling of isogenic KRAS wild type and mutant NSCLC cells in vitro and in vivo. Sci Rep. 2016 22;6:28398.

77. Hammond DE, Mageean CJ, Rusilowicz EV, Wickenden JA, Clague MJ, Prior IA. Differential reprogramming of isogenic colorectal cancer cells by distinct activating KRAS mutations. J Proteome Res. 2015 Mar 6;14(3):1535–46.

78. Zahari MS, Wu X, Blair BG, Pinto SM, Nirujogi RS, Jelinek CA, et al. Activating Mutations in PIK3CA Lead to Widespread Modulation of the Tyrosine Phosphoproteome. J Proteome Res. 2015 Sep 4;14(9):3882–91.

79. Wu X, Renuse S, Sahasrabuddhe NA, Zahari MS, Chaerkady R, Kim M-S, et al. Activation of diverse signalling pathways by oncogenic PIK3CA mutations. Nat Commun. 2014 Sep 23;5:4961.

80. Olsen JV, de Godoy LMF, Li G, Macek B, Mortensen P, Pesch R, et al. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. Mol Cell Proteomics MCP. 2005 Dec;4(12):2010–21.

81. Taus T, Köcher T, Pichler P, Paschke C, Schmidt A, Henrich C, et al. Universal and confident phosphorylation site localization using phosphoRS. J Proteome Res. 2011 Dec 2;10(12):5354–62.

82. Käll L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nat Methods. 2007 Nov;4(11):923–5.

83. Harsha HC, Pandey A. Phosphoproteomics in cancer. Mol Oncol. 2010 Dec;4(6):482–95.

84. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics MCP. 2002 May;1(5):376–86.

85. Mitchell CJ, Kim M-S, Na CH, Pandey A. PyQuant: A Versatile Framework for Analysis of Quantitative Mass Spectrometry Data. Mol Cell Proteomics MCP. 2016;15(8):2829–38.

86. Mahajan K, Mahajan NP. ACK1/TNK2 tyrosine kinase: molecular signaling and evolving role in cancers. Oncogene. 2015 Aug 6;34(32):4162–7.

87. Mahajan NP, Coppola D, Kim J, Lawrence HR, Lawrence NJ, Mahajan K. Blockade of ACK1/TNK2 To Squelch the Survival of Prostate Cancer Stem-like Cells. Sci Rep. 2018 Jan 31;8(1):1954.

88. Shinmura K, Kiyose S, Nagura K, Igarashi H, Inoue Y, Nakamura S, et al. TNK2 gene amplification is a novel predictor of a poor prognosis in patients with gastric cancer. J Surg Oncol. 2014 Mar;109(3):189–97.

89. Chan W, Sit S-T, Manser E. The Cdc42-associated kinase ACK1 is not autoinhibited but requires Src for activation. Biochem J. 2011 Apr 15;435(2):355–64.

90. Jang H, Banerjee A, Chavan T, Gaponenko V, Nussinov R. Flexible-body motions of calmodulin and the farnesylated hypervariable region yield a high-affinity interaction enabling K-Ras4B membrane extraction. J Biol Chem. 2017 28;292(30):12544–59.

91. Nussinov R, Wang G, Tsai C-J, Jang H, Lu S, Banerjee A, et al. Calmodulin and PI3K Signaling in KRAS Cancers. Trends Cancer. 2017 Mar;3(3):214–24.

92. Ye X, Chan KC, Waters AM, Bess M, Harned A, Wei B-R, et al. Comparative proteomics of a model MCF10A-KRasG12V cell line reveals a distinct molecular signature of the KRasG12V cell surface. Oncotarget. 2016 Nov 24;

93. Wu X-R, He X-S, Chen Y-F, Yuan R-X, Zeng Y, Lian L, et al. High expression of CD73 as a poor prognostic biomarker in human colorectal cancer. J Surg Oncol. 2012 Aug 1;106(2):130–7.

94. Wu R, Chen Y, Li F, Li W, Zhou H, Yang Y, et al. Effects of CD73 on human colorectal cancer cell growth in vivo and in vitro. Oncol Rep. 2016 Mar;35(3):1750–6.

95. Padanad MS, Konstantinidou G, Venkateswaran N, Melegari M, Rindhe S, Mitsche M, et al. Fatty Acid Oxidation Mediated by Acyl-CoA Synthetase Long Chain 3 Is Required for Mutant KRAS Lung Tumorigenesis. Cell Rep. 2016 09;16(6):1614–28.

96. Gouw AM, Eberlin LS, Margulis K, Sullivan DK, Toal GG, Tong L, et al. Oncogene KRAS activates fatty acid synthase, resulting in specific ERK and lipid signatures associated with lung adenocarcinoma. Proc Natl Acad Sci U S A. 2017 Apr 25;114(17):4300–5.

97. Aten TM, Redmond MM, Weaver SO, Love CC, Joy RM, Lapp AS, et al. Tyrosine phosphorylation of the orphan receptor ESDN/DCBLD2 serves as a scaffold for the signaling adaptor CrkL. FEBS Lett. 2013 Aug 2;587(15):2313–8.

98. Zhao R, Zhao ZJ. Dissecting the interaction of SHP-2 with PZR, an immunoglobulin family protein containing immunoreceptor tyrosine-based inhibitory motifs. J Biol Chem. 2000 Feb 25;275(8):5453–9.

99. Zhao R, Guerrah A, Tang H, Zhao ZJ. Cell surface glycoprotein PZR is a major mediator of concanavalin A-induced cell signaling. J Biol Chem. 2002 Mar 8;277(10):7882–8.

100. Zhao R, Zhao ZJ. Identification of a variant form of PZR lacking immunoreceptor tyrosine-based inhibitory motifs. Biochem Biophys Res Commun. 2003 Apr 18;303(4):1028–33.

101. Beigbeder A, Chartier FJM, Bisson N. MPZL1 forms a signalling complex with GRB2 adaptor and PTPN11 phosphatase in HER2-positive breast cancer cells. Sci Rep. 2017 Sep 14;7(1):11514.

102. Jia D, Jing Y, Zhang Z, Liu L, Ding J, Zhao F, et al. Amplification of MPZL1/PZR promotes tumor cell migration through Src-mediated phosphorylation of cortactin in hepatocellular carcinoma. Cell Res. 2014 Feb;24(2):204–17.

103. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature. 2009 Nov 5;462(7269):108–12.

104. Singh A, Greninger P, Rhodes D, Koopman L, Violette S, Bardeesy N, et al. A gene expression signature associated with "K-Ras addiction" reveals regulators of EMT and tumor cell survival. Cancer Cell. 2009 Jun 2;15(6):489–500.

105. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics MCP. 2002 May;1(5):376–86.

106. Kim DI, Cutler JA, Na CH, Reckel S, Renuse S, Madugundu AK, et al. BioSITe - a method for direct detection and quantitation of site-specific biotinylation. J Proteome Res. 2017 Dec 18;

107. Ori A, Free P, Courty J, Wilkinson MC, Fernig DG. Identification of heparin-binding sites in proteins by selective labeling. Mol Cell Proteomics MCP. 2009 Oct;8(10):2256–65.

108. Mitchell CJ, Kim M-S, Na CH, Pandey A. PyQuant: A Versatile Framework for Analysis of Quantitative Mass Spectrometry Data. Mol Cell Proteomics MCP. 2016;15(8):2829–38.

109. Prieto-Echagüe V, Gucwa A, Brown DA, Miller WT. Regulation of Ack1 localization and activity by the amino-terminal SAM domain. BMC Biochem. 2010 Oct 27;11:42.

110. Galisteo ML, Yang Y, Ureña J, Schlessinger J. Activation of the nonreceptor protein tyrosine kinase Ack by multiple extracellular stimuli. Proc Natl Acad Sci U S A. 2006 Jun 27;103(26):9796–801.

111. Chan W, Sit S-T, Manser E. The Cdc42-associated kinase ACK1 is not autoinhibited but requires Src for activation. Biochem J. 2011 Apr 15;435(2):355–64.

112. Kelley LC, Weed SA. Cortactin is a substrate of activated Cdc42-associated kinase 1 (ACK1) during ligand-induced epidermal growth factor receptor downregulation. PloS One. 2012;7(8):e44363.

113. Teo M, Tan L, Lim L, Manser E. The tyrosine kinase ACK1 associates with clathrin-coated vesicles through a binding motif shared by arrestin and other adaptors. J Biol Chem. 2001 May 25;276(21):18392–8.

114. Mahendrarajah N, Borisova ME, Reichardt S, Godmann M, Sellmer A, Mahboobi S, et al. HSP90 is necessary for the ACK1-dependent phosphorylation of STAT1 and STAT3. Cell Signal. 2017 Nov;39:9–17.

115. Kaizuka T, Hara T, Oshiro N, Kikkawa U, Yonezawa K, Takehana K, et al. Tti1 and Tel2 are critical factors in mammalian target of rapamycin complex assembly. J Biol Chem. 2010 Jun 25;285(26):20109–16.

116. Prieto-Echagüe V, Gucwa A, Craddock BP, Brown DA, Miller WT. Cancer-associated mutations activate the nonreceptor tyrosine kinase Ack1. J Biol Chem. 2010 Apr 2;285(14):10605–15.

117. Yeow-Fong L, Lim L, Manser E. SNX9 as an adaptor for linking synaptojanin-1 to the Cdc42 effector ACK1. FEBS Lett. 2005 Sep 12;579(22):5040–8.

118. Yokoyama N, Miller WT. Biochemical properties of the Cdc42-associated tyrosine kinase ACK1. Substrate specificity, authphosphorylation, and interaction with Hck. J Biol Chem. 2003 Nov 28;278(48):47713–23.

119. Pao-Chun L, Chan PM, Chan W, Manser E. Cytoplasmic ACK1 interaction with multiple receptor tyrosine kinases is mediated by Grb2: an analysis of ACK1 effects on Axl signaling. J Biol Chem. 2009 Dec 11;284(50):34954–63.

120. Jones S, Cunningham DL, Rappoport JZ, Heath JK. The non-receptor tyrosine kinase Ack1 regulates the fate of activated EGFR by inducing trafficking to the p62/NBR1 pre-autophagosome. J Cell Sci. 2014 Mar 1;127(Pt 5):994–1006.

121. Chua BT, Lim SJ, Tham SC, Poh WJ, Ullrich A. Somatic mutation in the ACK1 ubiquitin association domain enhances oncogenic signaling through EGFR regulation in renal cancer derived cells. Mol Oncol. 2010 Aug;4(4):323–34.

122. Sowa ME, Bennett EJ, Gygi SP, Harper JW. Defining the human deubiquitinating enzyme interaction landscape. Cell. 2009 Jul 23;138(2):389–403.

123. Hein MY, Hubner NC, Poser I, Cox J, Nagaraj N, Toyoda Y, et al. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell. 2015 Oct 22;163(3):712–23.

124. Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, et al. Architecture of the human interactome defines protein communities and disease networks. Nature. 2017 May 25;545(7655):505–9.

125. Lobingier BT, Hüttenhain R, Eichel K, Miller KB, Ting AY, von Zastrow M, et al. An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. Cell. 2017 Apr 6;169(2):350-360.e12.

126. Paek J, Kalocsay M, Staus DP, Wingler L, Pascolutti R, Paulo JA, et al. Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling. Cell. 2017 Apr 6;169(2):338-349.e11.

127. St-Denis N, Gupta GD, Lin ZY, Gonzalez-Badillo B, Veri AO, Knight JDR, et al. Phenotypic and Interaction Profiling of the Human Phosphatases Identifies Diverse Mitotic Regulators. Cell Rep. 2016 22;17(9):2488–501.

128. Mackmull M-T, Klaus B, Heinze I, Chokkalingam M, Beyer A, Russell RB, et al. Landscape of nuclear transport receptor cargo specificity. Mol Syst Biol. 2017 Dec 18;13(12):962.

129. Dong J-M, Tay FP-L, Swa HL-F, Gunaratne J, Leung T, Burke B, et al. Proximity biotinylation provides insight into the molecular composition of focal adhesions at the nanometer scale. Sci Signal. 2016 14;9(432):rs4.

130. Haugsten EM, Sørensen V, Kunova Bosakova M, de Souza GA, Krejci P, Wiedlocha A, et al. Proximity Labeling Reveals Molecular Determinants of FGFR4 Endosomal Transport. J Proteome Res. 2016 07;15(10):3841–55.

131. Kostas M, Haugsten EM, Zhen Y, Sørensen V, Szybowska P, Fiorito E, et al. Protein Tyrosine Phosphatase Receptor Type G (PTPRG) Controls Fibroblast Growth Factor Receptor (FGFR) 1 Activity and Influences Sensitivity to FGFR Kinase Inhibitors. Mol Cell Proteomics MCP. 2018 May;17(5):850–70.

132. Kim BR, Coyaud E, Laurent EMN, St-Germain J, Van de Laar E, Tsao M-S, et al. Identification of the SOX2 Interactome by BioID Reveals EP300 as a Mediator of SOX2-dependent Squamous Differentiation and Lung Squamous Cell Carcinoma Growth. Mol Cell Proteomics MCP. 2017 Oct;16(10):1864–88.

133. Redwine WB, DeSantis ME, Hollyer I, Htet ZM, Tran PT, Swanson SK, et al. The human cytoplasmic dynein interactome reveals novel activators of motility. eLife. 2017 18;6.

CURRICULUM VITAE

RAIHA TAHIR

2766 Mill Flume Drive, rtahir@jhmi.edu
Midlothian, Virginia, 23112 (804) 245-6688

EDUCATION

Ph.D.	2014-2018	Biochemistry, Cellular and Molecular Biology Program Johns Hopkins University School of Medicine Baltimore, MD Mentor: Akhilesh Pandey, M.D., Ph.D Thesis Focus: Mutation-specific and common phosphotyrosine signatures of *KRAS* G12D and G13D alleles *Anticipated graduation August 1st, 2018*
B.S.	2010-2014	Bioinformatics Virginia Commonwealth University Richmond, VA

RESEARCH EXPERIENCE

Summer Intern	2013-2014	Laboratory of Sean Taverna, Ph.D (Center for Epigenetics) Johns Hopkins University School of Medicine Baltimore, MD Project title: "Investigating the protein-protein interactions within the yeast histone acetyltransferase complex NuA3."
Researcher	2011-2013	Laboratory of Amanda Dickinson, Ph.D (Dept. of Biology) and Sarah Elsea, Ph.D (Dept. of Human and Molecular Genetics) Virginia Commonwealth University, Richmond, VA Project title: "Role of Retinoic acid induced-1 (Rai1) in regulating craniofacial and brain development in *Xenopus*."

PUBLICATIONS

Tahir R, Renuse S, Madugundu A, Cutler JA, Nirujogi RS, Na CH, Wu X, Pandey A, (2018). Mutation-specific and common phosphotyrosine signatures of KRAS G12D and G13D alleles. *Manuscript under review*

Tahir R, Madugundu A, Udainiya S, Cutler JA, Renuse S, Mitchell C, Wu X, Pandey A, (2018). Proximity-dependent biotinylation to elucidate the interactome of TNK2 non-receptor tyrosine kinase. *Manuscript under review*

Kim DI,* Cutler JA,* Na CH, Reckel S, Renuse S, Madugundu A, **Tahir R**, Goldschmidt HL, Reddy K, Huganir R, Wu X, Zachara NE, Hantschel O, Pandey A, (2017). BioSITe: A method for direct detection of site-specific biotinylation. *Journal of Proteome Research.* PMID: 29249144

Cutler JA, **Tahir R**, Sreenivasamurthy SK, Mitchell C, Renuse S, Nirujogi RS, Patil AH, Heydarian M, Wong X, Wu X, Huang TC, Kim MS, Reddy K, Pandey A, (2017). Differential signaling through p190 and p210 BCR-ABL fusion proteins revealed by interactome and phosphoproteome analysis. *Leukemia.* PMID: 28210003.

Tahir R, Kennedy A, Elsea SH, and Dickinson AJ. (2014) Retinoic acid induced-1 (Rai1) regulates craniofacial and brain development in Xenopus. *Mechanisms of development* PMID: 24878353

Mullegama SV, Pugliesi L, Burns B, Shah Z, **Tahir R**, Gu Y, Nelson DL, Elsea SH. (2014) MBD5 haploinsufficiency is associated with sleep disturbance and disrupts circadian pathways common to Smith-Magenis and fragile X syndromes. *European Journal of Human Genetics.* PMID: 25271084

MENTORSHIP/TEACHING EXPERIENCE

2013 Teaching Assistant
Introduction to Bioinformatics, Virginia Commonwealth University
Assisted students in completing programming assignments and project.

PRESENTATIONS

Tahir, R, Wu X, Zahari, Pandey, A (2017) The non-receptor tyrosine kinase TNK2/ACK1 is a novel therapeutic target in triple negative breast cancer. **Poster Presentation.** April 21, 2017. 10th Annual Breast Cancer Research Retreat, Baltimore, MD

Cutler, JA, **Tahir, R**, Kostadinov, R, Sekhar, R, Huang, TC, Brown, P, Heydarian, M, Pandey, A, Reddy, K.(2016) Proteomic/Transcriptomic Signatures of Infant MLL-r Rearranged B-ALL at Diagnosis and Relapse Reveal Lineage Plasticity and Diagnostic Heterogeneity **Poster Presentation.** December 1-4, 2016 ASH Annual Meeting & exposition. 128 (22), 2697-2697